

A Phonetic and Phonological Investigation of North American English (NAE) Segments
in the Interlanguage Grammar of a Native Speaker of German (SHG)

by

Lisa Süßenbach
B.A., Georg-August-Universität Göttingen, 2013

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Master of Arts

in the Department of Linguistics

© Lisa Süßenbach, 2018
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy
or other means, without the permission of the author.

Supervisory Committee

A Phonetic and Phonological Investigation of North American English (NAE) Segments
in the Interlanguage of a Native Speaker of German (SHG)

by

Lisa Süßenbach

Bachelor of Arts, Georg-August-Universität Göttingen, 2013

Supervisory Committee

Dr. John Archibald, Department of Linguistics, University of Victoria
Supervisor

Dr. Sonya Bird, Department of Linguistics, University of Victoria
Departmental Member

Abstract

Supervisory Committee

Dr. John Archibald, Department of Linguistics

Supervisor

Dr. Sonya Bird, Department of Linguistics

Departmental Member

This thesis investigates the L2 English pronunciation of a native speaker of German who has lived in western Canada for 25 years. The goal of the study was understand the defining features of his accent, to determine what factors contributed to his accent, and to characterize his interlanguage grammar. There are two opposing theories about L2 speakers' linguistic competence, encoded in what is called their *interlanguage grammar*: 1) The L2 speaker has several heterogeneous grammars at their disposal depending on discourse type (the socio-/psycholinguistic theory), so variable task performance is indicative of variable competence, and 2) variable task performance exists but it not indicative of variable competence. Instead, competence is a stable, homogenous system and it is performance that is variable (the generative/rationalist theory). This thesis discusses the concepts of variable competence in light of the L2 English pronunciation investigated.

The subject's pronunciation of a variety of speech sounds of North American English was tested in three production tasks with differing formality levels: wordlist, sentences, and a semi-spontaneous interview. Additionally, in a qualitative element of this study, extra-linguistic factors like motivation, attitude, aptitude, identity, and personality of the L2 speaker were investigated to determine how they contribute to L2 accented speech production. These were reported through an interview with the subject and a self-assessment of his L2 pronunciation proficiency. Finally, through native speaker judgments (NSJs), it was assessed whether foreign accentedness in the L2 interferes with intelligibility and comprehensibility. Production data from the three tasks was auditorily and acoustically analyzed to understand the contribution of various intra-linguistic factors to speech production: task type, orthography, cognate status, syllable context, stress, and phonetic environment. This thesis also investigated the validity of predictions made by the Speech Learning Model (Flege, 1995) about the ease of phonetic acquisition of L2 sounds.

The findings of this investigative study indicated that the L2 learner has a homogenous interlanguage grammar that is not responsible to variable competences despite variable task type performance. They showed that all variable performance in production could be attributed to intra-linguistic factors that influence performance, but do not alter the mental representation the subject has of these L2 sounds. Additionally, the findings showed that the Speech Learning Model does not accurately predict the ease or difficulty of acquisition of L2 speech sounds. Furthermore, the findings indicated that mispronunciation of individual speech sounds resulting in accentedness does not hinder effective communication in the L2, nor does accented speech production reflect an impoverished L2 interlanguage grammar. It further revealed that the subject was aware of his interlanguage grammar differing from that of native speakers of English. Findings

from the qualitative interview study indicated that the subject makes use of his accent as an identity marker to reflect his cultural attachment to his home country Germany.

Table of Contents

Supervisory Committee	ii
Abstract	iii
Table of Contents	v
List of Tables	vii
List of Figures	viii
Acknowledgments	ix
Dedication	xii
Chapter 1 Introduction	1
Chapter 2 Literature Review	8
2.1 Definition of Key Terms and Technical Terms	9
2.2 Foreign Accent in SLA	10
2.2.1 Definition of Foreign Accent and Measurement	10
2.2.2 Pedagogy's Focus on Intelligibility and Comprehensibility	12
2.3 The Interlanguage Grammar	15
2.3.1 Competence vs. Performance	18
2.4 Intra- and Extra-Linguistic Factors	22
2.4.1 The Speech Learning Model (SLM) (Flege, 1995) (Part I)	32
2.5 The Phonetic Inventories of NAE and SHG	37
2.5.1 Substitution of NAE /æ/ with SHG /ɛ/	40
2.5.2 Substitution of NAE /u:/1 with SHG /u:/2	40
2.5.3 Voicing contrast of obstruents appearing in coda position	41
2.5.4 Aspiration of voiceless stops in coda position in German	43
2.5.5 Substitution of NAE Interdentals /θ/ and /ð/ with SHG /s/ and /z/	45
2.5.1 Substitution of dark [ɫ] with light [l]	46
2.5.2 Substitution of NAE /w/ with SHG /v/	47
2.5.3 NAE /ɹ/ and SHG /ʁ/	49
2.5.4 Speech Learning Model (Part II): Predictions for NAE Segments	50
Chapter 3 Methodology	53
3.1 Participants	53
3.2 Instruments and Stimuli	53
3.3 Procedure	58
3.4 Data Analysis	61
Chapter 4 Results	64
4.1 Results on Substitution of /æ/	64
4.2 Substitution of NAE /ʌ/ with SHG /a/	69
4.3 On-target /ɒ/	72
4.4 Substitution of NAE /u/ with SHG /u/	75
4.5 Devoicing of /b, d, g, v, ð, z, ʒ, dʒ/ in coda position	78
4.6 Aspiration of /p, t, k/ in coda position	85
4.7 Interdentals	87
4.8 Epenthetic [t] / Non-reduction of /t/	90
4.9 Substitution of NAE [ɫ] with SHG [l]	96
4.10 Substitution of /w/ with /v/	99
4.11 Prevocalic /ɹ/	108
4.12 Substitution of /ɹ/ in Postvocalic Position	111

4.13	Results of the Native Speaker Judgments	119
Chapter 5	Discussion.....	125
5.1	Overall Accuracy	125
5.2	Vowels.....	127
5.3	Consonants	128
5.4	Intra-Linguistic and Extra-Linguistic Factors	131
5.4.1	Task Type	133
5.4.2	Orthography	134
5.4.3	Cognates	136
5.4.4	Phonetic Environment.....	138
5.4.5	Syllable Context.....	139
5.4.6	Stress	140
5.4.7	Self-Assessment and interview: extra-linguistic factors	141
5.5	M.S.' s Interlanguage Grammar	147
5.6	Speech Learning Model	159
5.7	Global Accentedness, Comprehensibility, and Intelligibility	166
5.8	Limitations and Future Directions	167
5.9	Concluding Remarks.....	169
Chapter 6	Reference List.....	172
Chapter 7	Appendix	182
Appendix A	182
Appendix B	184
Appendix C	186
Appendix D	187
Appendix E	188

List of Tables

Table 2.1 Definition of Key and Technical Terms	9
Table 2.2 Consonant chart of NAE and SHG	37
Table 2.3 SLM Classifications of NAE and SHG sounds	51
Table 4.1 M.S.'s productions of /æ/	65
Table 4.2 M.S.'s productions of /ʌ/	70
Table 4.3 M.S.'s productions of /ɒ/	72
Table 4.4 M.S.'s productions of /u/	76
Table 4.5 M.S.'s productions of underlyingly voiced coda obstruents	80
Table 4.6 Syllabic Environment Conditioning M.S.'s Obstruent Coda Voicing	83
Table 4.7 M.S.'s productions of Coda Voiceless Stops	86
Table 4.8 M.S.s productions of Interdentals	88
Table 4.9 Syllabic Context Conditioning M.S.'s Interdental Productions ("O-T" = On-target)	89
Table 4.10 M.S.'s productions of /CtC/ clusters	91
Table 4.11 Selection of tokens from the wordlist containing epenthetic [t] for M.S.	91
Table 4.12 Examples of /CtC/ in NAE by C.E. and M.S.	92
Table 4.13 M.S.'s productions of /l/	97
Table 4.14 M.S.'s productions of /l/ by Syllabic context	99
Table 4.15 M.S.'s productions of /w/	100
Table 4.16 Comparative Features of NAE and SHG Labial sounds	103
Table 4.17 NAE and SHG <w> Cognates	104
Table 4.18 Vowel Environment by Task conditioning /w/ productions	105
Table 4.19 Summary of Vowel Environment conditioning /w/ productions	106
Table 4.20 M.S.'s productions of /v/	107
Table 4.21 M.S.'s productions of Prevocalic /ɪ/	109
Table 4.22 Comparative Feature chart for /ɪ/	111
Table 4.23 M.S.'s productions of Postvocalic /ɪ/	112
Table 4.24 Summary of Native Speaker Judgements	122
Table 5.1 Summary of M.S.'s Results split by vowel vs. consonant	126
Table 5.2 M.S.'s Phonetic/Phonemic Vowel differences	128
Table 5.3 M.S.'s Phonetic/Phonemic Consonant differences	130
Table 5.4 M.S.'s Overall Accuracy of segment productions in descending order	131
Table 5.5 Factors affecting M.S.'s Productions	132
Table 5.6 Extra-linguistic factors affecting M.S.'s L2 speech	144
Table 5.7 Self-Assessment of Problematic NAE Segments	147
Table 5.8 The State of Segments in M.S.'s Interlanguage Grammar	149
Table 5.9 Accuracy of SLM Predictions	161

List of Figures

Figure 2.1 Spectrogram of the vowel [ʌ] showing formants (F1 and F2).....	9
Figure 2.2 Two Differing Views of Interlanguage Grammar.....	22
Figure 2.3 Classifications of L2 Sounds.....	34
Figure 2.4 NAE and SHG vowels of interest in comparison.....	39
Figure 4.1 M.S.'s Acoustic productions of /æ/ compared to NAE standards.....	67
Figure 4.2 M.S.'s Acoustic productions of /ʌ/.....	71
Figure 4.3 M.S.'s Acoustic productions of /ɒ/.....	74
Figure 4.4 M.S.'s Acoustic productions of /u/.....	77
Figure 5.1 Influence of the L1 and L2 on the Interlanguage Grammar	148

Acknowledgments

First and foremost, I will “turn into William Faulkner, king of long sentences in novels” to thank my supervisor Dr. John Archibald for his immeasurable support and encouragement throughout the pursuit of this thesis, even though I’ll try to keep it as short as possible, hoping that I will be cut some slack, for I told myself not to spend more than two sentences on every person, although this is rather difficult taking into account how fantastic of a supervisor you have been, John. PERIOD. It has been a rocky road, and I must say, I don’t recall a single time you did not go out of your way to provide feedback, read chapters, share smart thoughts, all topped with that sense of humour of yours – most importantly, though, you kept me grounded when I wanted to do 100 things at once. You are truly the best supervisor anyone could ever ask for and I cannot thank you enough. I feel honoured and humbled to have been your grad student and truly learned so much.

I would also like to thank my second member, Dr. Sonya Bird, whose meticulous feedback and advice helped bringing this thesis to fruition. Thanks for the talks, the books, and the ideas you’ve shared with me, including your honest feedback. I wanted to let you know that I appreciated every comment you made that contributed to improving this thesis – positive or negative, time-consuming or easy fix.

Furthermore, I would like to acknowledge Dr. Catherine Caws for her support, commitment, and compassion as my outside member. A special ‘thank you’ goes to Dr. Margaret Cameron, Associate Dean of Research. You’ve been wonderful in every single way. No further words needed – just gratitude. I would also like to thank my colleagues from Germanic and Slavic Studies “downstairs” for the opportunities, support, and

friendship you've given me throughout the years: Helga, Ulf, Charlotte, Elena, Matt, Kat, and Olga.

Secondly – and equally important – I would like to thank my family. I know me being fully committed to academia hasn't always been easy for you, but I can assure you that this has been the right path for me. Thank you for all those years of financial support, emotional support, and accepting and respecting my choices and decisions: Oma, Mama, Papa, Ela, Lena, Bine, Jule, Martin, and my wonderful cousins (whom I cannot enumerate). I love all of you dearly.

Finally, even though all of this would have *technically* been possible without you, it would have been utterly terrible without you, my really good friends from 'good old Germany': Sandy, Schoqui, Sonni, Carmen, Hyri, Laura, Julia, Alina, Ben, Gina, "Mama Birgit", and everyone I forgot. Blame my head, not my heart. All of you have contributed to me keeping my sanity with your love and support, even though we live miles apart now. Special thank-yous go to my former professors from the University of Göttingen who have inspired and encouraged me beyond the level of the BA. You are amazing scholars and researchers with wonderful brains and hearts: Dr. Hildegard Farke, Prof. Dr. Daniel Stein, Prof. Dr. Oliver Traxel, Diana Rosenhagen, and last but not least: Dr. Harald Kittel.

I also want to thank my friends from Victoria, including my colleagues from the Linguistics Department, most notably Silas, Dustin, Janet, Kyra, Yiran, Jie, Marjolein, and Ildara – thank you for your friendship and companionship. You sure have been of great help to keep me going and pushing through.

Ulf, what can I say? You are truly the best friend anyone could ever ask for, and if I had the chance to give someone a Superman Award, it would be you. Tess – I'm in complete awe about who you are, how much you know, and how incredibly kind you've been to me. You are a great friend to have, in all respects. Steph, you've been hilarious and a dedicated friend throughout the years, your sense of humour is truly unique, and I miss you dearly. Finally, I'd like to thank Mari; you're the perfect friend – just when you have to get away from it all and need good advice, great coffee, and a yoga lesson. You're the best mom I know. Annye, merci pour tout ce-que tu as fait pour moi: tu es vraiment spéciale. And of course – last but not least – Megan: it's too bad you left Victoria, but I miss you and the endless laughs and time spent.

I'd also like to take the time to apologize to everyone I forgot to mention.

Dedication

To Marieke, my sweet baby cousin: You are the living proof why language is such a wonderful tool. One day I will apologize to you for taking joy in your frustration of not being understood, but you got me into phonetics and phonology with this one sentence:

“Habe du hauch Huhe han?” Why, yes, I do have shoes on. You were merely a little child.

Chapter 1 Introduction

When describing foreign accented speech, native speaker listeners draw on perceptual (and acoustic) cues of how vowels and consonants are produced. In the past decades, research has provided a relatively rich picture of German-accented speech production of English segments (Hanulikova & Weber, 2012 ; Flege & Bohn, 1990 ; Lombardi, 2003 ; Charles-Luce, 1985 ; Smith, Hayes-Harb, Bruss & Harker, 2009). The existing studies were mainly quantitative ones that concerned themselves with production data by large numbers of participants and were therefore focused on a specific segment of English that was deemed problematic for native speakers of Standard High German (henceforth: SHG). Among those segments are ‘classics’ that English language learners coming from different native languages experience difficulties with, such as interdental fricatives (Bien, Hanulikova, Weber & Zwitserlood, 2016 ; Hanulikova & Weber, 2012) or obstruents in coda position, which are devoiced in many languages (German included) but not in English (Charles-Luce, 1985 ; Smith, Hayes-Harb, Bruss & Harker, 2009). Up to this day, and to my best knowledge, there has not been any case study of an L1 German L2 English speaker that covers an extensive number of North American English (henceforth: NAE) segments that can all pose potential difficulties to native speakers of German.

The segments that may pose difficulties are considered to be those that have anecdotally been reported and/or addressed in monographs on German and English phonetics, based on measures of contrastive analysis (König & Gast, 2009 ; Grantham O’Brien & Fagan, 2016 ; Swan & Smith, 2001 ; Hall, 2003). The *Contrastive Analysis Hypothesis* (Lado, 1957) compares the sounds of two languages and predicts production

errors based off of differences. However, we now know that comparing two sound systems of relatively related languages, which German and English are, can only provide limited insight into a system of L2 grammar that an individual has acquired, and that simply performing a contrastive analysis of any two sound systems cannot fully predict speech production errors. Not all production errors can be attributed to transfer of the L1 to the L2 (Archibald & Libben, 1995), and so not all production errors can be predicted by contrastive analysis. Instead, models like the *Speech Learning Model* (Flege, 1995) allow us to make predictions from the premise that perception of new sounds precedes production, and so it is sounds that are similar which will be harder to acquire (because they will be harder to perceive as different from sounds of the L1). We may find that there are sounds that differ completely from those we know, and that these are easy to acquire, while there are also sounds that are similar to the ones we know and are more difficult to acquire. In addition, not every native speaker of German with English as a second language will exhibit the same kinds of errors, because the speech production of L2 learners is variable and affected by a variety of factors, both intra-linguistic (orthography, task type, cognate status, stress, phonetic environment, and syllable context) and extra-linguistic in nature and learner-specific (motivation, identity, personality, anxiety, attitude, and aptitude) (Archibald, 1998). Since these extra-linguistic factors are idiosyncratic, it can also be useful to know how the speaker perceives their own L2 production, providing insight into variables that the researcher cannot glean from quantitative measurements and analysis alone.

Previous studies (Tarone, 1987; Ellis, 1997; Dickerson, 1974; Major, 1991) have explored the L2 grammar the language learner has acquired (the *interlanguage grammar*)

through the notion of ‘variable competence’, meaning that language learners use several L2 grammars depending on different contexts and discourses. This variation is often investigated across different task types and held to stand for different contexts/discourses. There is a widely held belief that an L2 learner’s production differs depending on formality of the task, which may support the idea that an individual has access to several interlanguage grammars.

The present study seeks to investigate the *interlanguage grammar* of a 50-year-old male with L1 German and L2 English and seeks to understand how English segments are manifested in his interlanguage after having lived in Canada for 25 years of his life. Taking the intra- and extra-linguistic factors just defined into account and combining these variables, the goal of the study is to paint a picture of the subject’s interlanguage grammar of English. In order to achieve this goal, the following research questions were generated:

1. What does the production of NAE segments reveal about the interlanguage grammar in pronunciation in a native speaker of SHG who has lived in Canada for 25 years?
2. What intra- and extra-linguistic factors can account for accented L2 speech production?
3. Does production performance differ across task types (wordlist, sentences, semi-spontaneous production) and is there such a thing as ‘variable competence’?

4. What statements can a foreign language learner make about his or her own accent and phonetic and phonemic production errors?

5. Does SHG-accented NAE speech production contribute to decreased intelligibility and comprehensibility ratings by native speaker listener judges (NSJs) of NAE?

To understand the speaker's interlanguage grammar, several methods are utilized. Accuracy scores for a number of NAE target segments (/æ/, /ʌ/, /ɒ/, /u/, voicing contrast in coda of /b, d, g, v, ð, z, ʒ, dʒ/, aspiration in coda of /p, t, k/, /θ, ð/, epenthesis of [t] / non-reduction of /t/ in consonant clusters, [ɸ], /w/, and prevocalic and postvocalic /ɪ/) will be employed to obtain an overview of how well these segments have made themselves a home in the subject's speech production. Because L2 learners have variation in how they acquire different speech sounds from the L2, they may find some L2 sounds more difficult to acquire than others. The difficulty or ease with which a sound may be acquired can be predicted through the *Speech Learning Model* (Flege, 1995) based on the *newness* or *similarity* L2 sounds have to L1 sounds. Based on Flege's model, predictions are made about which of the North American English sounds listed above should be easy to acquire, and which ones should pose difficulties to a native speaker of German.

These specific segments are examined through the lens of intra-linguistic factors (given above) that contribute to non-native speech production on the segmental level which can influence performance. A major factor that is said to influence production performance in an L2 is task type formality. Therefore, several task types were employed to elicit speech production from the subject in three different contexts: wordlist (most

formal), sentence reading (formal), and semi-spontaneous speech (least formal). This allowed for a direct assessment of the role that discourse/context may play in variable competence/performance.

Moreover, this case study allowed for the inclusion of an interview which served as a self-assessment to obtain insight into the speaker's perception of the constitution of his own interlanguage grammar. One has to determine what escapes the radar of being produced in a L2-like manner and what the L2 speaker himself perceives as erroneous about his own speech production. The self-assessment and interview can also provide information on extra-linguistic effects on accent, such as the subject's identity, motivation, and attitude, and why segmental errors and foreign accent are not necessarily viewed as negative from the perspective of the L2 speaker (Derwing, 2003 ; Derwing and Rossiter, 2002; Achirri, 2017).

While segmental errors certainly contribute to foreign accented speech production, in order to judge how or if this actually hinders communication, we need native speaker listeners to quantify foreign accent. Doing so draws on three concepts: 'intelligibility' (how much of an utterance is understood by a listener), 'comprehensibility' (the effort required to understand the utterance), and 'accentedness' (how different the speaker sounds from the listener) (van Maastricht, Krahmer & Swerts, 2016). This thesis will explore how segmental errors contribute to perceived non-native accent by native listeners of English, and whether these segmental errors are grave in that they hinder intelligibility or comprehensibility.

Using the methods described above to answer the research questions set out earlier, this thesis makes a novel contribution as a mixed-methods approach to exploring,

in-depth, the interlanguage grammar of a German L2 English speaker. Combined with the native speaker judgments of the L2 English speaker's perceived global accentedness and comprehensibility, this thesis is not only able to demonstrate that he is a proficient and comprehensible speaker of NAE, but also to challenge the notion of 'variable competence' depending on task type. In fact, in bringing together linguistic and extra-linguistic factors of speech production, a picture of a fixed system of the interlanguage grammar of his L2 emerges that suggests that there is no 'variable competence' throughout task types, despite occasional variable production performance which could be attributed to the various linguistic factors. Furthermore, the thesis demonstrates that the Speech Learning Model does not reliably predict which sounds of the L2 are easy or difficult to acquire and produce.

The rest of the thesis proceeds as follows: Chapter 2 begins with definitions of some key and technical terms used throughout this thesis (2.1), followed by a review of the literature on foreign accent and second language acquisition (2.2), and an overview of the theory of interlanguage grammar and competence vs. performance (2.3). Then it discusses the various intra- and extra-linguistic factors that may have an effect on M.S.'s speech (2.4) as well as the Speech Learning Model (SLM), and how the model defines and predicts the acquisition of *new* and *similar* sounds (2.4.1). Following this, there is an overview of NAE sounds predicted by the literature to be problematic for German L2 English learners along with descriptions of the differences and similarities of these sounds with SHG sounds, and finally the SLM predictions for these sounds (2.5).

Chapter 3 introduces the methods used: a discussion of the participants and how they were selected (3.1), the instruments and stimuli used (wordlists, sentences,

interview, questionnaire, native speaker judges) and how they were designed and chosen (3.2 & 3.3), and the methods used to evaluate and analyze the data collected via these instruments and stimuli (transcriptions, acoustic analysis, statistical tests) (3.4). Chapter 4 presents the results of these data collection procedures, starting first with production data: segments (going from vowels to consonants) (4.1 to 4.12), and then moving to how native listeners perceive the production data (through Native Listener Judgments) (4.13).

Chapter 5 discusses the results in the light of the factors introduced in Chapter 2. A distinction in the behaviour and analysis of vowels and consonants is made (5.2 & 5.3), followed by a detailed discussion of the various factors influencing M.S.'s interlanguage: task type, cognates, orthography, phonetic environment, syllable context, stress, and extra-linguistic factors (5.4). Then the insights gained into M.S.'s interlanguage grammar from the investigation of these factors are discussed and the notion of 'variable competence' is refuted (5.5). Finally, there is an evaluation of the predictions of the Speech Learning Model and how the SLM failed to accurately predict many of the results (5.6), plus a general discussion of global accentedness, comprehensibility, and intelligibility in M.S.'s speech overall (5.7), followed by a brief discussion on the limitations of the current work, and areas for future research (5.8).

Chapter 2 Literature Review

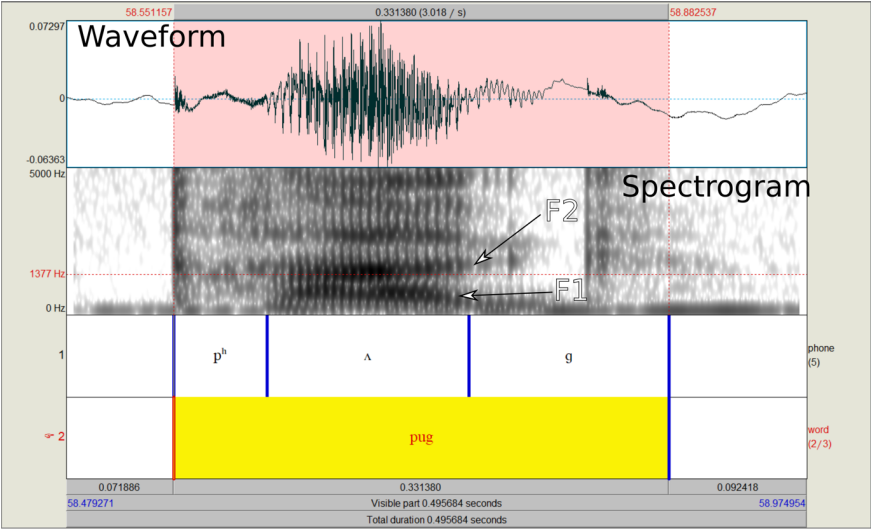
In this chapter I will first lay out technical terms as they will appear throughout this thesis (2.1). This will be followed by an introduction of concepts that pertain to the research questions on foreign accent in second language acquisition and how such accent can be measured (2.2). Due to these factors influencing L2 speech, we can assume that production performance is not always stable and that it can vary depending on discourse type, e.g., production accuracy can be linked to task type formality. Following the description of those factors that contribute to variable production performance by L2 speakers, a section on the interlanguage system that L2 learners develop over the course of acquisition of a second language will be introduced (2.3). Production of L2 speech is subject to several factors of intra-linguistic and extra-linguistic nature that can influence L2, all of which will be addressed in this chapter (2.4). I will explain what *interlanguage grammar* is and provide an account of different notions of its constitution: does an L2 speaker employ several grammatical systems in their L2 depending on task formality (variable competence) or is there just one underlying interlanguage system and variable task production can be attributed to performance only (variable performance)?

As this study focuses on the acquisition of NAE speech sounds in a native speaker of SHG, the phonetic inventories of NAE and SHG will be compared and pronunciation difficulties that may arise for native speakers of SHG acquiring NAE speech sounds will be discussed (2.5). In order to account for differing degrees of difficulty of the acquisition of NAE sounds by SHG native speakers, I will adopt the notion of the Speech Learning Model by Flege (1995).

2.1 Definition of Key Terms and Technical Terms

Throughout this thesis, the technical terms and key terms listed in the table below will be used when reporting results and findings and discussing their implications. These terms are presented in order of relevance to the structure of the thesis, with the first four terms pertaining directly to aspects of the results (formants and co-articulation) or the quantification of results (accuracy and markedness) and the last two terms referring to theoretical concepts brought up in the discussion to frame and interpret the results.

Table 2.1 Definition of Key and Technical Terms

Term	Definition
formants (F1, F2)	<p>Bandwidths of frequencies (measured in Hertz) that have greater intensity (loudness) than other frequencies in a sound. Formants can be found in sonorous sounds, like vowels, and provide information about a speaker's vocal tract and its configuration. Each of the formants correlates with an articulatory feature of a vowel: F1 = vowel height F2 = vowel frontness/backness (Ladefoged, 2006)</p> 
co-articulation	<p>Overlap between two speech sounds, such as nasalization of a vowel preceding a nasal stop, e.g., /pæn/, <i>pan</i>, as [pæ̃n]</p>

accuracy	Although it is difficult to set a threshold for accuracy (in this study: on-target productions) that would clearly indicate a sound as 'acquired' vs. 'not acquired', general practice adopts a threshold of >80% accuracy in on-target productions to consider an element of speech as 'acquired'. (Eckman, 1991)
typological markedness	certain linguistic elements are less likely to appear across the languages of the world, as they are distinctively defined, making them 'marked', whereas the more common linguistic structures across languages are considered more 'general' and therefore 'unmarked.' (Eckman, 1977)
performance and competence	Chomsky (1965) uses the terms <i>competence</i> and <i>performance</i> and differentiates the two by stating that <i>competence</i> refers to a mental property to form grammatical speech, whereas <i>performance</i> only refers to the production of utterances. Thus, competence is the actual 'knowledge' of a language, whereas performance entails what we do with it, and is affected by the situation in which we are actually speaking.
interlanguage (IL) grammar	A linguistic system an L2 speaker of a given language has established which is comprised of L1 and L2 elements as well as novel elements (Selinker, 1972)

2.2 Foreign Accent in SLA

In this section of the literature review I will introduce general concepts of Second Language Acquisition (henceforth: SLA) research on foreign accent, how it can be described and measured, and the factors that contribute to the nature of a foreign accent.

2.2.1 Definition of Foreign Accent and Measurement

“By definition, accent encompasses all the layers of phonetic and prosodic precision required to convey and negotiate meaning” (Moyer, 2013). All native speakers of English can usually distinguish a foreign ESL accent from an English L1 variant accent. This seems to be universal to humans in general, who can usually tell if someone has a foreign accent. Interestingly enough, though, we even have the ability to identify nonnative speakers of languages we do not even know (Major, 2007). This is not necessarily detected by listening to a single phoneme in isolation – although it can be – but by listening to accented speech productions containing substitutions, deletions, or insertions that violate pronunciation rules or the phonotactics of English. Such variations are usually detected by a native speaker of English, for example. While lexis, morphology,

and syntax are domains in which nonnative speakers can achieve high proficiency, L2 pronunciation remains one of the most difficult obstacles to overcome in second language acquisition (Hulstijn, 2015).

As studies of the past five decades have shown, very few late L2 learners achieve nativelike pronunciation, which is believed to be linked to “maturational effects of the brain” (Lenneberg, 1967). Thus, it is believed that with progressing age, nativelike attainment of the L2 becomes increasingly difficult, if not entirely impossible for late learners of a foreign language (Abrahamsson & Hyltenstam, 2009). Not just age effects are believed to play a role in maintaining a foreign accent. More importantly, the onset of meaningful L2 exposure and the age of onset (AoO) of acquisition is believed to be a “robust predictor of success in second language acquisition” (Granema & Long, 2012). Previous studies on foreign accent have also studied the importance of length of residence (LOR) (Best & Tyler, 2007; Flege, 1995; Flege, Schirru, & MacKay, 2003; Major, 2008) and have produced conflicting findings about the duration of exposure to the L2 in an L2 speaking environment affecting the L2 pronunciation. In fact, LOR resulting in higher target-like performance in the second language is believed to depend on many more cognitive and psychological individual variables of the learner, such as motivation, attitude, aptitude, anxiety, and personality (Archibald & Libben, 1995 ; Saito & Brajot, 2013 ; Moyer, 2013), as will be further explained in section 2.3 on extra-linguistic factors.

Despite all these factors playing a role in L2 pronunciation attainment, some studies report “exceptional” learners that do manage to acquire an L2 in a nativelike manner (Bongaerts, Summeren, Planken, & Schils 1997 ; Colantoni & Steele, 2006 ;

Coppieters, 1987 ; White & Genesee, 1996 ; Birdsong, 1992). However, it remains realistic to believe that these learners constitute the minority of language learners and that very few late learners will be able to pass themselves off as native speakers.

2.2.2 Pedagogy's Focus on Intelligibility and Comprehensibility

Within the field of pedagogy, it used to be popular to attempt accent reduction – or even complete accent elimination – in nonnative speakers; a goal that is deemed unattainable for (most) nonnative speakers of a language (Munro & Derwing, 1995 ; 1999 ; Murphy, 2014). Especially for late learners, attempting foreign accent reduction has been deemed a fruitless endeavour as there is no evidence to believe that at one point the second language learner will achieve nativelike pronunciation (Abrahamsson & Hyltenstam, 2009).

English is a global lingua franca, and has twice as many L2 speakers as L1 speakers overall (Simons & Fennig, 2018 ; Thomson & Derwing, 2015), and so there has been a shift that turned focus away from the reduction of foreign accent towards the concept of being able to effectively communicate in the L2. For example, drawing on the goal of native-like pronunciation attainment, researchers like Derwing and Munro (1995) argue that it is only realistic to assume that the primary goal of the foreign language learner should be to be intelligible and comprehensible in L2 communication. Upon introducing the terms *intelligible* and *comprehensible*, one has to establish what these concepts refer to. These two concepts relate to another concept, namely that of *accentedness*. *Intelligibility*, *comprehensibility*, and *accentedness* capture different

aspects of L1 listeners' perception of L2 speech (Munro & Derwing, 1995 ; 1999). How then can these terms be defined?

Accentedness

Accentedness merely denotes a difference in the quality of the pronunciation between speaker A (nonnative speaker) and listener B (native speaker), which is established by the latter. The listener perceives “how closely the pronunciation of an utterance approaches that of a native speaker” (Kennedy & Trofimovich, 2008, p.461).

Intelligibility

The concept of intelligibility refers to “the extent to which a given utterance is understood by a listener”.

Comprehensibility

Comprehensibility refers to the degree of ease or effort that is required by the listener to understand the speaker.

We can now establish that the concepts of accentedness, intelligibility, and comprehensibility relate to the perception of the native listener, which is a reliable means to foreign accent detection, as they have an “... [accurate] perception of deviations from a pronunciation norm that [they] attribute [...] to the talker not speaking the target language natively” (McCullough, 2013). With this in mind, any native speaker of English qualifies as a judge to rate ESL speech production, for example. Native listeners can perceive errors on the segmental (individual sounds) and the suprasegmental levels (stress, intonation, rhythm) (Kennedy & Trofimovich, 2008 ; Munro & Derwing, 1995 ; Murphy, 2014). However, research has shown that phonetic errors and even phonemic

errors usually do not affect pronunciation intelligibility and comprehensibility (Thomson & Derwing, 2015).

How then are accentedness, intelligibility, and comprehensibility measured?

Intelligibility is usually measured by orthographic transcription the native speaker listeners provide while listening to a speech utterance (Kennedy & Trofimovich, 2008). The orthographic transcription of an utterance is then measured against a transcription of the recorded utterance. Every time there is a discrepancy, it is counted as an error that decreases intelligibility (Munro & Derwing, 1995 ; Derwing & Munro, 1997).

To measure comprehensibility, the native speaker (NS) listeners usually provide a ‘goodness’ rating on a Likert scale that can range from 1 to virtually any number the researcher deems reasonable. Most researchers employ a Likert scale that ranges from 1-9 (Munro & Derwing, 1995; 1999), the number increasing with effort the NS listener requires to understand the utterance: 1 would correspond to “easy to understand” on the continuum, whereas 9 would be “extremely difficult to understand”, which the listener judge can indicate on a spreadsheet. Likewise, for accentedness, the NS listener listens to the same utterances and assesses how “accented” the speech production is. This is done by the same measures of employing a 1-9 Likert scale, 1 would equal “no foreign accent”, whereas 9 would be “very thick foreign accent”. Previous studies have found that increased accentedness of an L2 speaker does not necessarily result in decreased intelligibility. Thus, even heavily accented speech can still be perfectly intelligible. On the other hand, there is a negative correlation between comprehensibility and accentedness rating, with low perceived accentedness of an utterance resulting in elevated comprehensibility scores (Munro & Derwing, 1995; 1999 ; Murphy, 2014).

Bringing together these three concepts, we see their relationship and find that accented speech production can be perfectly intelligible and comprehensible (Derwing, 2010) and, consequently, any L2 speaker with a significant accent can still be considered proficient and/or fluent in their L2.

However, one has to point out that these judgments on accentedness, intelligibility, and comprehensibility can be primarily used to assess the speaker's L2 proficiency on a global level. While phonemic and phonetic errors can be detected through native speaker listener transcription of utterances, the mispronunciation of a single segment will not reveal much about intelligibility or comprehensibility, as speakers ultimately have to use connected speech in real life situations, which is the only situation in which measuring for intelligibility and comprehensibility seems appropriate (Thomson & Derwing, 2015). The native speaker judgements included in this study will serve as a means to establish what the connections are, in connected speech, between segmental production errors and intelligibility and comprehensibility, demonstrating that merely having an accent cannot prevent effective communication in the L2 (Thomson & Derwing, 2015 ; Munro & Derwing, 1995).

2.3 The Interlanguage Grammar

Drawing on the notion of foreign accent and its constitution, we have established that most (late) foreign language learners will exhibit a foreign accent in their L2. This foreign accent may be perceived due to the L2 learner exhibiting segmental productions of L2 speech sounds that are not those of the target language. In this section on *interlanguage grammar*, we will learn how foreign accent relates to the state of an L2

learner's system of linguistic competence, with special focus on the domain of pronunciation.

When analyzing how second language learners pronounce L2 speech sounds, one can often describe and measure what they do, e.g., Germans typically substitute [s] for /θ/, so NAE *think* sounds like NAE *sink*, but it is seldom explained *why* they do it or that this substitution may be due to a certain system that within itself is “productive and rule-governed” (Archibald & Libben, 1995). While interlanguage encompasses all domains of SLA lexis, grammar, syntax, and pronunciation, the notions of interlanguage will be explained here as they relate to L2 pronunciation. The interlanguage can be regarded as a stage of mental representation that is located between the grammar of the L1 and the grammar of the L2. It is not a deficient version of the L2, as behaviourist approaches claimed for a long time in line with *Contrastive Analysis Hypothesis* and the notion of the negative “interference” of the L1 with the L2 (Skinner, 1957; Lado, 1957). This model predicts that a given learner will transfer structures from his L1 speech to the L2 in those areas in which the L2 differs. Thus, all pronunciation errors can be attributed to *L1 transfer*. However, when looking more closely at pronunciation errors made by L2 learners, it is safe to say that not all of them can be attributed to L1 transfer and that there is more to the story of how L2 speech sounds are acquired.

The interlanguage grammar of an individual L2 learner does **not** progress from an initial L1-like form to a nativelike L2-like form over time, with all L2 pronunciation forms increasing and slowly replacing L1 forms. Rather, the L2 learner acquires a system, a competence for his or her L2 that consists of elements from the L1 (those elements that are transferred), the L2 (those elements that have been acquired) and an “in-between

area” which we call the *interlanguage* that exhibits elements that are neither coming from exclusively the L1 or exclusively the L2. Rather, both the L1 and the L2 constitute the interlanguage which consists of patterns described by Tarone (1987) below, such as overgeneralization, approximation, and avoidance. Finally, an L2 learner’s interlanguage system is made up of elements of the L1, the L2, and the interlanguage (Archibald & Libben, 1995). As classified by Selinker’s notion of *fossilization* (1972), we know that very few L2 learners reach native-like proficiency, as explained in Section 2.2.1. When we speak of fossilization, we acknowledge that for individual areas of L2 acquisition, a final stage before L2-nativeness has been reached, in which the proficiency no longer progresses. This stage or state, of course, is what we find in most non-native speakers of a language. And for each learner of a L2 this may be different. Not every language learner reaches native-like attainment; there are points along the way where progress stops for a learner. The state of a given interlanguage depends on when in the process of acquisition a given L2 element became fossilized. The three major processes Tarone (1987) describes are illustrated below in reference to SHG and NAE:

1. **Overgeneralization.** A native speaker of SHG produces NAE [w] and starts using this sound with [v] interchangeably, not having acquired the /w ~ v/ contrast, and so they start using [w] randomly for English /w/ and /v/ (Chamson, 2016; discussed in more detail in Section 2.5).

2. **Approximation** can be explained at the example of a native speaker of SHG producing a sound that is neither found in their L1, nor their L2. With the example of NAE /w/ again, which is absent from SHG, once the learner has understood how /w/

differs from /v/, they might start to produce an in-between sound like [v̥], which has the quality of a glide, just like /w/, but is produced without lip rounding, just like /v/.

3. **Avoidance** would be a strategy the L2 learner uses to circumlocute production of a difficult sound. For example, the German native speaker may find it difficult to produce NAE interdental fricatives in codas, like in *both*. So instead, he might escape by paraphrasing or just using “*the two of them*”, to avoid a difficult sound in a syllable position where it is the most difficult to produce.

2.3.1 Competence vs. Performance

These processes outlined by Tarone are key properties of the interlanguage and can never be purely attributed to L1 transfer exclusively, or acquisition of an L2 element. Instead, drawing on the examples given above on overgeneralization, approximations and avoidance, we find elements that are neither home to the L1 nor the L2.

We can think of the interlanguage grammar as a separate and self-contained system, different from both L1 and L2. When we think of a system, we think of an organized, homogenous set of principles. However, here I discuss two competing views of what makes up this separate and self-contained Interlanguage system. After going over the distinction of ‘competence’ and ‘performance’, I detail the idea that there is such a thing as ‘variable competence’, a notion that is motivated by differences in task type performance by L2 speakers. This approach is motivated by, and draws from, psycho- and socio-linguistic traditions. I then discuss the competing generativist-rationalist view that evidence for variable competence is left wanting, and the examples of ‘variable competence’ are actually just ‘variable performance’.

Firstly, what is the difference between ‘competence’ and ‘performance’? This distinction is a Chomskyan notion. Competence is invariable, which means that the L2 learner has **one** homogenous state of his or her interlanguage grammar and therefore tacit (unconscious) linguistic knowledge (generative/rationalist approach). This competence reflects the state of the learner’s phonological mind. Performance is the actual expression and use of language in the world between individuals and is subject to a variety of conditions ranging from one’s physical ability to speak to tiredness, to the constraints of memory, or endurance. Among L1 speakers, when speech errors are made, they are considered to be due to these performance factors and not deficient knowledge of the language as part of linguistic competence.

Researchers like Dickerson (1975) and Yule & Tarone (1991) suggest that L2 competence is subject to task formality, assuming that the L2 learner’s performance may differ from one task to another, and that this reflects a difference in competence. Tarone (1987), Ellis (1984), and Major (2001) take research that has shown that task type production can vary from the most formal (e.g. wordlist) to the least formal (e.g. spontaneous speech production) as evidence for an underlying ‘variable competence’ that is responsible to the formality of the discourse situation of an L2 speaker. The fact that the same L2 speaker can exhibit different production accuracy depending on the formality of the task has served as the motivation for the notion that the system of the L2 speaker itself is **not** fixed, but variable (Ellis, 1984 ; Tarone, 1987). Tarone and Dickerson disagree on one point: Dickerson claims that production is most accurate in formal tasks because of monitoring and where attention can be focused on one token, like a wordlist, whereas Tarone claims the opposite, namely that informal task productions are

the most accurate because of the triggering of L1 structures (Silveira, 2007). Studies conducted over the years show inconclusive results (Major, 1994; Lin 2001; Silveira, 2007). But while task type effects will be further discussed in section 2.4 (on Task Type), the question that arises is whether differing performance on task types suggests that interlanguage grammar also has ‘variable competence’ (sociolinguistic/psycholinguistic approach).

Ellis (1990) claims that “the competence of the learner is much more variable than that of the native speaker, for the simple reason that interlanguage systems are not as strongly fixed or established as fully-formed natural languages, which makes them more permeable to new forms that enter the system” (p.387). Therefore, he suggests that competence is subject to change, and that new forms of grammar make it into the system of the language learner, where they co-exist with another old form. To illustrate this, one can refer to the production of the NAE interdental /θ/, for example, by native speakers of German. Ellis’s hypothesis is that at one point, the native speaker of German will use SHG /s/ for NAE /θ/. As the learning process progresses, NAE /θ/ will be acquired but phonemically co-exist with /s/, and the learner will use the two sounds interchangeably, for example. Thus, the representation keeps changing over time (Brown, Malmkjaer, & Williams, 1996). Major (2001) notes that performance differences are a result of variable competence, thus, different grammars that were acquired at some point become activated depending on the discourse situation. This can also be seen in the socio-linguistic variation among native speakers, who vary in production when talking to their peers (less formal speech register) and when employing a more formal speech register. According to Major (2001), an 80-year-old man showed phonetic variation of English /ɔ/ and /ɑ/ in

spontaneous speech, but showed a near merger in minimal pairs, which is typical of younger speakers. Thus, he is claimed to show variable competence depending on discourse type.

In contrast to those arguing that competence can be variable, Gregg (1990) supports the Chomskyan distinction of competence and performance and argues that the L2 learner does not apply a variable rule, which means that “in given circumstances a given person has a given probability of using a given variant” (Jordan, 2004). He acknowledges that there is “varying use of forms”, but claims that this does not tell us anything about the underlying linguistic knowledge/competence of the language learner. Instead he argues that linguistic factors influence production performance (Jordan, 2004). Gregg (1990) claims that Tarone, Ellis, and Major merely describe phenomena of variable *performance* rather than competence and mistake this for variable competence. For example, when Tarone (1987) describes that task production differs from the most formal to the least formal type, she describes the differing production performance and attributes this to ‘variable *competence*’ without providing any evidence about the state of the underlying competence (the IL grammar) of the L2 speaker or demonstrating that the phonological mind of the L2 learner changes depending on task type due to him or her employing different grammars for different discourse situations

According to Gregg, we are dealing with one homogenous competence that, regardless of task type, does not change. In other words, the phonological mind of the language learner is not varying from task type to task type, and while the output changes, it does not mean that a subject has native-like competence for an L2 element in **one** task

and non-nativelike competence in another. The competence, the phonological mind of the L2 speaker does not change from one task to the other, only his or her performance does.

Concluding, we have two opposing views: the argument that there is ‘variable competence’ (Tarone, 1987 ; Ellis, 1984) as the different levels of mastery of certain speech styles show, as opposed to the argument that there is ‘one competence’ and that stylistic variation is a result of variable performance (Major, 2001 ; Gregg, 1990). The approaches are summarized in the figure below:

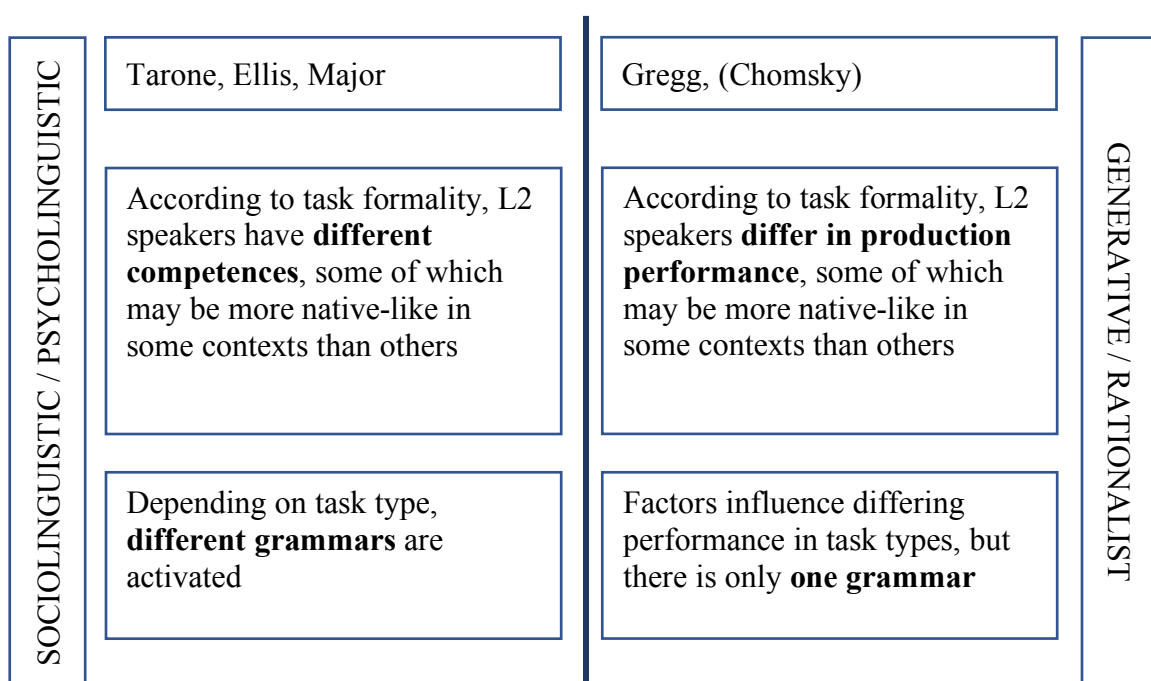


Figure 2.2 Two Differing Views of Interlanguage Grammar

2.4 Intra- and Extra-Linguistic Factors

In this section, different variables that can influence L2 pronunciation will be discussed. These factors were deemed crucial to examine in this thesis, as they provide insight into the composition of the interlanguage grammar. As explained in section 2.3, Tarone (1987) and Ellis (1984) claim that differing task type performance is evidence that L2

learners have access to different grammars. However, other intra-linguistic factors can also account for speakers *performing* differently across tasks, and this need not mean that the underlying grammatical representation changes. These factors can be divided into two sets: intra-linguistic factors (task type, orthography, cognate status, syllable context, phonetic environment, and stress) and extra-linguistic ones which are of cognitive or psychological nature of the L2 learner. While intra-linguistic factors universally apply to any language learner, extra-linguistic factors (motivation, attitude, anxiety, personality type, and identity) are learner specific in that they pertain to individual psychological and cognitive traits.

Task Type

It is widely assumed that performance of the L2 speaker is variable depending on context and discourse (Hansen Edwards, 2008). Several studies have shown that more formal tasks, such as word reading tasks, can lead to more monitoring of speech output, whereas in a spontaneous production task, one would pay less attention to one's output, and therefore do less monitoring of the speech (Hansen Edwards, 2008). Now, what does this mean for predicted accuracy across tasks? Previous studies reported that production is most accurate in formal tasks (wordlist) and then declines as one progresses to the least formal task (spontaneous production) (Gatbonton, 1975, Dickerson, 1975). Thus, there is widely believed to be a correlation between speech monitoring, which is more likely to occur in formal tasks, and production accuracy.

On the other hand, there is also a substantial amount of research that reports the opposite pattern: L2 learners are more accurate in production in the less formal task (Moyer, 2004; Sato, 1985; Tarone & Parrish, 1988). These authors were hesitant to claim

that task formality and accuracy are correlated, suggesting other factors may play an even more crucial role, such as phonological environment, and some extra-linguistic factors, like educational background or social class (Hansen Edwards, 2008).

While it is widely acknowledged that task performance *does* indeed vary depending on task type formality, a larger debate that follows the one above is what these differing performances of L2 speakers reveal about their actual underlying competence foreign language learners have in their L2. In fact, researchers like Tarone (1987) and Ellis (1984) take differing task type performance effects as evidence of **different competences**. Ergo, if production varies from task type to task type, it is due to the L2 speaker accessing different kinds of grammars or competences of their L2. Other researchers like Gregg (1990) refute the theory that variable task performance is evidence for L2 speakers employing different L2 grammars in each task, respectively. Gregg believes that task type variation can only account for variable production performance. He acknowledges that L2 speakers perform differently depending on discourse situation, but doubts that this can be attributed L2 speakers having access to multiple L2 grammars.

One of the aims of this study is to see whether different task types with different formality levels have any bearing of the production of M.S.'s NAE segments and whether those can be attributed to variable competence (Tarone, Ellis, & Major) or variable performance (Gregg).

Orthography

When a study involves different formality tasks, the more formal tasks (e.g. wordlists and sentences) will often involve orthographic input. A common intuition is that orthography affects the L2 speaker's pronunciation. However, we know that if the subject had been

exposed to L2 orthographic input, orthographic effects can also be found in spontaneous speaking tasks without orthographic input, as for learners of English in spontaneous speech production (Silveira, 2007). One of the explanations why (alphabetic/syllabic) orthography influences L2 production beyond the level of a formal reading task is that L2 learners (who are literate in their L1 and L2) have orthography in their mental representation of L2 lexical items (Archibald & Young-Scholten, 2000).

Archibald and Young-Scholten (2000) believe that language learners growing up being exposed to second language orthography from the beginning of first learning would have constant access to mental representations of the spelling of a word, which would influence their L2 production, even in a less formal task, for example, in an interview (Watkins, Rauber, & Baptista, 2009). Empirical studies that have investigated the relation of orthography and L2 production have found a correlation between the two, but whether orthographic knowledge helps or hinders is still up in the air (Escudero, Hayes-Harb, & Mitterer, 2008 ; Escudero & Wanrooij, 2010).

English is generally a language in which the orthography is incongruent with its phonology, meaning that single graphemes do not directly correspond to single phonemes. A good example to demonstrate this would be the different pronunciations for *though*, *tough*, and *ought*, which share the same spelling, but are all pronounced differently: [ðʊʊ], [tʌf], [ɒt]. German, on the other hand, has a much more transparent orthography, in which phoneme-grapheme correspondence is much higher (Snowling & Hulme, 2013). For example, vowels are pronounced the same in *Altlasten*, *Kanone*, and *Apfel*: ['altlastən], [ka'no:nə], and ['apfəl]. Since knowledge of L1 orthography will influence the L2, one must consider orthographic effects that could influence formal tasks

as well as informal task productions. For this study, orthographic factors either from the L1 or the L2 could be assumed to play a role in the subject's production of NAE segments. For a native speaker of German, having such high grapheme-phoneme correspondence from in his L1, one can suspect that they will articulate English L2 sounds that are seemingly present according to orthography, but that are actually absent. E.g.: NAE *instruction* has a <t> grapheme, but there is no underlying /t/ sound. Similarly, NAE *though* has a <gh> grapheme, but there is no underlying sound present. Therefore, orthographic input, especially in formal tasks, can be misleading in that sounds are articulated due to a grapheme suggesting they should be. Referring back to the example of NAE *instruction* above, SHG has the same word coming from Latin in its inventory: *Instruktion* where <t> does represent a sound, so in in German, words ending in the morpheme {-tion} always have /t/: [ɪnstʁʊkt͡ʃjoːn]

Cognates

Another important intra-linguistic factor that influences L2 production is that of cognates. What are cognates? Previous research defines cognates as a pair of words in the L1 and L2 that have a shared etymology, roughly similar or identical meaning, and a shared phonological form (Szubko-Sitarek, 2014 ; Parkes & Cornell, 1996). For example: NAE *poodle* and SHG *Pudel* are etymologically related (i.e., they share a common origin in an earlier language), as are NAE *terrace* and SHG *Terrasse*. Thus, cognates can be triggered through orthographic resemblance, semantic overlap, and/or phonology (Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010). Dijkstra et al. (2010) conducted a lexical access study on English words that have a Dutch competitor in any of the above-mentioned domains. The results revealed that “orthographic and semantic overlap

facilitated lexical decision responses” (Dijkstra, et al., 2010). In fact, this definition of a cognate is a common way to identify and teach about cognates.

Carroll (1992) claims that the etymology of words is unrelated to cognate status in the mind of a learner. In other words, the learner does not know about a shared origin of words just because they are related. Instead, it is the formal representation, the surface structure of a word is the most decisive factor. For example, two words can be absolutely unrelated etymologically or semantically, such as NAE *bright* and SHG *breit* ‘wide’, NAE *gift* and SHG *Gift* ‘poison’, or NAE *eagle* and SHG *Igel* ‘hedgehog’, but the learner would use the formal representation of the word (auditory or spelling) and associate the two. For this reason, this thesis will focus just on the formal representation of cognates: what do they look like in spelling or sound like phonologically, and how does this influence the L2 production?

As far as pronunciation is concerned, formal (phonological) cognate status inhibits lexical retrieval. Thus, the phonology of the L1 is competing with the phonology of the L2 and “...the fact that this is only seen in L2 can be explained by the relative dominance of the L1 pronunciation over the L2 pronunciation of cognate words” (Midgley, Holcomb, & Grainger, 2011). According to the Homogeneity Hypothesis by Libben (2000), the lexical representations for the L1 and the L2 are the same in the mental lexicon in L2 learners (Dijkstra et al., 2010). These are accessed during lexical processing. So even in a spontaneous production task, cognate status can play a role for mental representation that the learner has access to. For example, take the word NAE *bear*. There is semantic overlap in both languages (SHG *Bär*) and the mental representation for *bear* would be the same in the L1 as in the L2, but the German L2

learners would access their L1 SHG *Bär* which would then influence their L2 production because of the cognate status and the phonological and therefore orthographical representation of the word NAE *bear* in the L2. It is important to note that cognate status and orthography are often confounded. Considering an example from the previous section on orthography, we can determine that cognate and orthographic effects are difficult to tease apart. A fitting example would be the words English and German share from Latin ending in <-tion>, like in NAE *instruction*/ SHG *Instruktion*. A <t> grapheme is present in both the English as well as the German term, but it is only pronounced as /t/ in German, not in English. So we see, because of their status as cognates, their etymological relation, and their formal orthographic presentation (both are almost identical in orthography), how cognate status and orthography often go hand in hand.

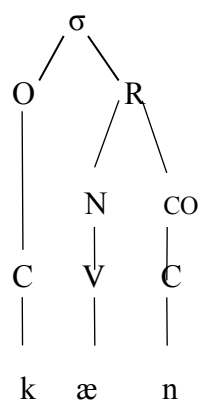
Phonetic Environment

Important to consider in the production of speech sounds is the surrounding phonetic environment. Specific parts of an utterance are subject to adjacent sounds' influence. This can make utterances of the same underlying phoneme sound different due to conditioning by surrounding sounds, an effect of what we call co-articulation (Crystal, 2011). One example that can show how surrounding sounds can influence the utterance of a given sound would be the nasalization of vowels in NAE. Vowels can change in quality when a nasal sound follows them, in words like *pan* [pãn] vs in *pat* [pæt] or *Canada* [kãnidə] vs *cat* [kæt]. Of course, this is due to the consonant /n/ being nasal, therefore the preceding vowel has a nasal quality to it. To provide another example – this time from German – we can look at the uvular fricative /χ/ like in the word *Bach* 'brook' [baχ] and its allophone the palatal fricative [ç] like in the word *ich* 'I' [ɪç]. These sounds are not

contrastive in German, because they are both underlyingly /χ/ (Kohler, 1999); each allophone is conditioned by the vowel environment in which it occurs. There is a systematic pattern that will have [χ] following all back vowels and [ç] following all front vowels, because back vowels would be articulated closer to uvula for [χ], and front vowels would be articulated closer to the hard palate like [ç]. This means that producing this allophonic variety in SHG is an effect of co-articulation. When considering the data production for M.S.'s productions, one can assume that such co-articulatory effects as observed in NAE and SHG can influence accuracy scores. These are not necessarily just conditioned by the L1, but these may be universal effects that can be found in L1 speakers too.

Syllable Context

Another factor we have to consider when analyzing L2 production is that of syllable context. There are three elements to a syllable (though not all need be present for an allowable syllable): an onset (C), a rhyme containing a nucleus (V), and a coda (C), like in the word *can*, for example:



The syllable position in which a segment appears can affect production. For example, NAE interdental /θ/ and /ð/ are relatively rare across languages, which makes them what we would call “more marked” (Simons & Fennig, 2018 ; Moran, McCloy, & Wright, 2014) . They are only found in 4% of all languages across the world. Being more marked often correlates with the degree of difficulty of attainment of these sounds. In contrast, the sound /m/ appears in 95% of all world languages (Moran, McCloy, & Wright, 2014). Eckman (1977) assumes that a sound of a foreign language is more difficult to learn if it is not present in the L1 *and* is also typologically less common across world languages. Therefore, for a native speaker of SHG, a sound like /θ/ or /ð/ would be hard to acquire, because a) It does not exist in SHG and b) it does not exist in many world languages. It is similar for syllables. Syllables of a CVC structure are more marked than syllables of a CV structure, and furthermore when the coda consonant is allowed it tends to contain only a certain subset of all the consonants allowed in onset position (Clements, 1990 ; Blevins, 1995 ; Zec, 1995). Bringing segmental and syllable markedness together, we can say that if a rare (marked) segment appears in coda position (also marked), this would be particularly difficult to produce. Therefore, we can assume that the position of a syllable conditions the production of sounds.

Stress

Stress correlates with vowel quality, and thus with vowel production. A language like English reduces vowels when they occur in unstressed position. For example: *communication* [k^həmjʊnɪ'k^heɪʃn] versus *commerce* ['k^hɒməs] bear different syllable stress, which leads to reduction of the underlying vowel /ɒ/ in *communication*, but not in *commerce* because the first syllable bears stress in the latter word, but not the former.

Thus, when we compare *communication* and *commerce* the syllable <com-> also has different duration in each of these words, respectively. [k^həm] in *communication* has a shorter duration than [ˈk^hɒm] in *commerce*. This means that when stress increases the salience of a particular syllable, production is affected. The stressed syllable can have greater duration, and other syllables that are not stressed can have reduced duration and altered vowel quality. Thus, for the L2 speaker, stressed syllables may offer a favourable environment for vowel production, as the conditions of stress (longer duration, more salience and less reduction of quality) increase the likelihood of on-target production.

It was for this reason that this study took into consideration stress of vowels. Because of the potential for increased production accuracy (and higher accuracy scores) that stressed vowels may exhibit, due to their salience addressed above.

Other Factors (Extralinguistic)

Undoubtedly, there are other factors which could be examined to analyze the segmental production of M.S.'s speech production, because we know that we cannot predict the same production output for each L2 learner of English with SHG as a native language. Indeed, if we were to contrast the sound system of NAE with SHG, as promoted in the CAH (Lado, 1957), and draw rigid conclusions that each native speaker of SHG makes the same production errors, we would not account for individual learner differences, such as cognitive and psychological factors. We cannot simply compare SHG to NAE and predict that each speaker of SHG will behave the same in NAE (Archibald & Libben, 1995). Extra-linguistic factors relevant to consider include: aptitude, attitude, anxiety, motivation, IQ, age, learning styles, gender, personality (Dörnyei & Ryan, 2015; Archibald & Libben, 1995; Moyer, 2013). We can safely assume that certain personality

types (e.g. introverted vs. extroverted) are more prone to anxiety when they have to converse in a foreign language. Further, we can acknowledge that some second language learners are exceptionally talented and have a “good ear”. Other language learners may be very interested in the culture in which they acquire a second language and will be more motivated to attempt more native-like pronunciation, as they have a positive attitude towards getting in touch with people who speak their first language (Dörnyei & Ryan, 2015). In contrast, some studies have shown that the maintenance of an L1 accent is deliberate and serves as an identity marker, which the L2 speaker wishes to keep to stay attached to their native language (Seidlhofer, 2001).

While diving into the realms of extra-linguistic factors in detail is not possible in this thesis, it is crucial to know that we cannot just predict what kind of accent a foreign language learner of a given language exhibits without taking a variety of the intra- and extra-linguistic factors discussed above into account. Those extra-linguistic factors pertaining to attitude, aptitude, identity, motivation, personality, and anxiety were anecdotally addressed in the interview and information on them is analyzed in chapters 4-5 to gain a fuller understanding of the participant’s accent.

2.4.1 The Speech Learning Model (SLM) (Flege, 1995) (Part I)

In Section 2.2, I talked about segmental articulation contributing to foreign accent. We could also establish that – for some L2 sounds – native-like proficiency may never be achieved, whereas others may be acquired by at least some individual learners (section 2.3 on Interlanguage). This suggests that the degree of difficulty of acquiring L2 sounds may not be the same for all sounds. In this section I detail a model that seeks to account

for why some of these segments are more or less likely to be acquired, the Speech Learning Model (Flege, 1995).

There are several theories that concern themselves with the acquisition of L2 sounds and the difficulties for nonnative speakers/listeners that may arise in either perception or production of these sounds. In the context of this study and its focus on L2 **production** I have decided to adopt Flege's Speech Learning Model (1995). The first initial model of how L2 speech sounds are acquired was Lado's *Contrastive Analysis Hypothesis* (1957), which compared the sound systems of two given languages (L1 and L2) and predicted production errors based on simply comparing their phonetic inventories. Lado concluded that if an L2 sound is absent from the L1, its acquisition will be difficult. On the other hand, if the sound in a given L2 is similar, it will be easy to acquire. While this initially seemed appealing and logical, this model implied that all speakers who share the same L1 make the same L1 transfer errors. *Contrastive Analysis Hypothesis* (henceforth: CAH) became subject to severe criticism in the 1970s and 1980s because it failed to account for individual learner performances and also claimed that production errors stem from "negative *L1 transfer*", which is the notion that the L1 will negatively interfere with the L2 if a given structure in the L2 differs from that of the L1, explained further in Section 2.3 above.

The method of comparing two sounds systems and pointing out the areas of difference leading to acquisition problems was eventually refuted, as learners have been

shown to have difficulties in areas in which the L1 and L2 differ, *as well* as in areas in which the L1 and L2 overlap (Richards & Sampson, 1974 ; Richards, 1974).

There was a paradigm shift in the 1990s from CAH, which brought a fundamentally different theory to the table: the Speech Learning Model (henceforth: SLM) by Flege (1995). According to Flege (1995), perception ability precedes production ability. Therefore, he argued that “difficulties in perception are responsible for difficulties in production” (Derwing & Munro, 2015, p. 68). Challenging the notion of CAH, the SLM suggests that it is the *similarity* of a given L2 sound to a familiar L1 sound that interferes with nativelike mastery and poses difficulties to the nonnative speaker, and not the *dissimilarity* (Bylund, Hyltenstam, & Abrahamsson, 2013). Therefore, accurate production of L2 sounds strongly correlates with the accurate perception of L2 sounds: because only if L2 sounds are perceived as distinct from L1 sounds can L2 speakers accurately produce them (Derwing & Munro, 2015). Flege bases his model on a trichotomous spectrum (including *identical*, *similar*, and *new* sounds) on which the L2 speaker either perceives an L2 sound as *identical* to an L1 sound or *new*. *Similar* sounds of an L2 have an L1 near-equivalent, but they, too, will be perceived as *identical*, which is why the spectrum reflects that only *new* sounds will be perceived as nothing L1-equivalent.

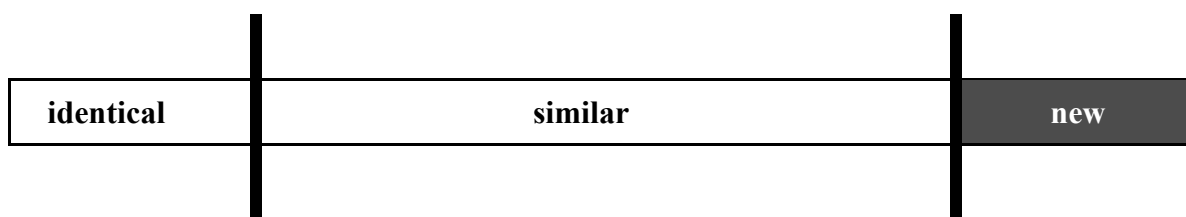


Figure 2.3 Classifications of L2 Sounds

According to Flege, *identical* sounds will pose no difficulties to the L2 learner, as there is perfect phonetic (and acoustic) overlap between an L1 and an L2 sound. They are in fact the same sounds. For example, NAE /s/ is identical to SHG /s/, therefore a native speaker of German would not experience difficulties perceiving and producing NAE [s]. Therefore, *identical* sounds of NAE and SHG will not be investigated and discussed in this thesis. The crucial components of the SLM lay in those sounds that are *similar* and *new*. As opposed to CAH, and at first counterintuitive, the SLM predicts that it is not *new* L2 sounds that will be difficult for the learner, but *similar* ones. It is important to know that Flege's "equivalence classification" is central to his theory (Flege, 1987 ; Hansen-Edwards & Zampini, 2008). The SLM suggests increased perception ability for nonnative speakers when the L2 sound is distinct from any L1 sound, which means that the more phonetically different any L2 sound is from the L1, the easier it is for the L2 learner to perceive and hear it as a *new* sound. Consequently, L2 sounds that are perceived as *similar* to L1 sounds by the L2 speaker are perceived as equivalents of sounds of the L1 and are therefore mapped onto an already familiar L1 sound and stored in the same phonetic category. "Equivalence classification will have the effect that the learner's perceptual norm for the L2 will merge with the norm that he has established previously for the equivalent in L1 [...]" (James & Leather, 1997). This can be demonstrated with the example of SHG /u:/ and NAE /u:/. Because the phonetic quality between these sounds is very similar, but not the same, the SHG speaker learner will perceive NAE [u:] to be an equivalent member of SHG /u:/. Thus, experience with an L1 sound, in this case SHG /u:/ will result in decreased ability to detect differences between that native phonetic category and the nonnative phonetic category of NAE /u:/ (Colantoni, Steele, & Escudero

2015). This can increase production difficulty, as the similarity of an L2 sound is perceived as identical to an L1 sound, when, in fact, they are different. While Flege assumes that L2 categories are formed in the same way native language categories are formed by children acquiring an L1, the phonetic categories of the L1 become increasingly fixed from childhood through adolescence. This means that the phonetic categories are so strongly established in the mind that a phonetically similar L2 sound will be mapped onto the closest L1 phonetic category, so formation of a new category for the L2 sound is blocked (Flege, 2016). However, for *new* L2 sounds, meaning those L2 sounds that are phonetically distant from any sound of L1, Flege predicts that they will be perceived as not L1 equivalent and will therefore be acquired more easily without interference from any L1 category (Bohn & Munro, 2007).

Older versions of the SLM often refer to the formal resemblance of IPA symbols for reference to define *new* and *similar* sounds (Flege, 2016). Classification on this basis was criticized due to its arbitrariness. Instead, it is now “the degree of *perceived* similarity or dissimilarity of the L2 sound by the L2 speaker/listener [that] determine[s] whether new L2 categories can be established” (Yang, 2016). Flege assumes an interplay of acoustic similarity and perceived similarity but does not give a specific or solid definition of what acoustic factors can correlate with perceptual factors, rather just that perceived similarity relates to acoustic similarity. Flege acknowledges that the boundary between *new* and *similar* sounds has not been established empirically by previous studies, but one can employ acoustic measures to determine closeness of an L2 and L1 sound.

I hypothesize that the sounds that differ in only one phonetic property (acoustic and articulatory feature) can be classified as *similar*, as the phonemic categories for each

of them overlap in NAE and SHG. On the other hand, an NAE sound would be classified as *new* if it differed from the closest SHG sound in more than one phonetic property (Flege, 2016). Each sound is investigated in turn in Section 2.5 below, and I make specific predictions based on these criteria for each.

2.5 The Phonetic Inventories of NAE and SHG

It can be established that the overwhelming majority of NAE and SHG consonants overlap, while their vowel inventories greatly differ. The consonant chart below illustrates the similarities and differences of the consonantal inventories of SHG and NAE:

Table 2.2 Consonant chart of NAE and SHG

Place Manner	bilabial	labiodental	interdental	alveolar	alveo- palatal	palatal	velar	uvular	glottal
stop	p b			t d			k g		ʔ
fricative		f v	θ ð	s z	ʃ ʒ	ç		χ	h
affricate		pf		ts	tʃ dʒ				
nasal	m			n			ŋ		
lateral				l			(ɮ)		
rhotic				ɹ				ʁ	
glide	w					j	(w)		

The black IPA symbols represent sounds that are present in both the inventories of SHG and NAE. Red IPA symbols represent sounds that are absent from SHG. Green IPA symbols represent sounds that are absent from NAE. For my thesis, and consequently for

this section, only the sounds highlighted in red, which are present in NAE but not in SHG, will be of interest.

Among those NAE segments of interest in this study are a few commonly reported to be “problematic” to the native speaker of German (König & Gast, 2009 ; Swan & Smith, 2001). For example, the NAE interdental fricatives /θ/ and /ð/ as well as the labiovelar glide /w/ are completely absent from the phonetic inventory of SHG. Moreover, the NAE alveolar liquid /ɹ/ and the velarized alveolar lateral [ɭ], for which SHG has a uvular fricative /ʁ/ and an alveolar lateral /l/, respectively.

In addition to segmental differences, SHG and NAE have different phonological processes affecting phoneme realization. Two processes will be considered here: a) final devoicing of obstruents in coda position and b) final aspiration of stops in coda position. These processes are illustrated below:

SHG has no voicing contrast of obstruents in coda position like NAE does where it contrasts voiceless and voiced obstruents:

/p, t, k, f, θ, s, ʃ, tʃ/ /b, d, g, v, ð, z, *¹(ʒ, dʒ)/

SHG only has voiceless or devoiced obstruents in syllable-final position. Conversely, SHG has aspirated voiceless stops in coda-position: [p^h, t^h, k^h], where NAE only has released or unreleased stops: [p, t, k] or [p̚, t̚, k̚].

As far as NAE vowels are concerned, German has a larger vowel inventory than English in general but lacks the unrounded low front vowel /æ/. Moreover, studies have found that while NAE and SHG both have phonemic /u/ in their respective vowel

¹ *Note that /ʒ/ and /dʒ/ exist in SHG in loanwords from French and English, but SHG itself only has /ʃ/ in its native inventory,

inventory, the NAE /u/ differs from that of NAE in phonetic quality. Kent and Read (2002) summarized 10 different studies of NAE vowel productions and provided their means, whereas Pätzold & Simpson (1997) did the same for SHG vowels, which have been adapted from the *Kiel Corpus of Read Speech*. These are plotted in Figure 2.4 below.

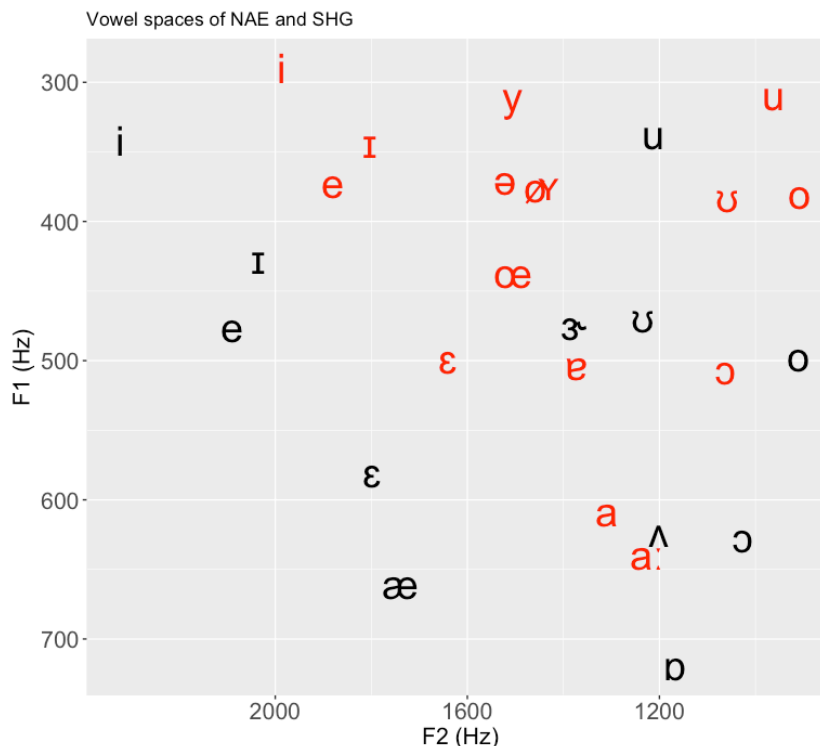


Figure 2.4 NAE and SHG vowels of interest in comparison

The above chart shows the comparative acoustic vowel spaces of male speakers of NAE (in black) and SHG (in red). Males speakers' means were used, as the subject of the study is a male.

In the following sections, the NAE sounds of interest (consonants and vowels) that are absent from SHG are presented with their substitution patterns. When a sound is absent from the L1 inventory, the second language learner may turn to substitutions,

deletions or epentheses. This overview is based on the substitution patterns of NAE sounds that have been observed and reported in existing literature.

2.5.1 Substitution of NAE /æ/ with SHG /ɛ/

As seen in the chart above, and as several studies have shown, SHG lacks the NAE low, unrounded front vowel /æ/, which is why native speakers of SHG substitute it with /ɛ/, as anecdotally described by Swan & Smith (2001). /æ/ is located between SHG /a:/ and SHG /ɛ/. In line with this description are the findings of previous studies that report that native speakers of SHG do, in fact, replace English /æ/ with the mid, unrounded front vowel [ɛ] (Bohn & Flege, 1990 ; 1992) and favour this substitution over that of SHG [a:]. A study by Bohn and Flege (in James & Leather, 1997) suggests that while [æ] and [a:] are acoustically not any more different than [æ] and [ɛ], [æ] is consistently replaced with the front vowel [ɛ]. While it cannot be explained why the substitution [ɛ] is favoured for /æ/, research has replicated these results time and time again, and it has been reported to happen as well with Dutch (Collins & Mees, 1984), Swedish, French and Italian speakers (Swan & Smith, 2001), all of which also have the low central /a/ and low mid front vowel [ɛ]. Bohn and Flege (1990) classified NAE /æ/ as a *new* sound, because it falls into a phonetic space in which German does not have a phonetic category (James & Leather, 1997). Therefore, its acquisition is predicted to be *easy* for SHG speakers.

2.5.2 Substitution of NAE /u:/1 with SHG /u:/2 ²

The SHG round back vowel /u:/ is produced as more back than its English counterpart (Grantham O'Brien & Fagan, 2016 ; Hall, 2003), and is also more peripheral, meaning,

² Unfortunately, despite the different quality of NAE /u:/ and SHG /u:/, conventional phonetic transcription does not differentiate symbolically between the vowels.

that SHG /u:/ gravitates towards the “edge” of the vowel space, as demonstrated in the above chart, whereas NAE /u:/ is much more centralized (Steinlen, 2005). The English /u:/ is also significantly less rounded than the German /u:/, so German *Mut* ‘courage’ and English ‘moot’ are not phonetically identical. To the native German speaker, however, the NAE /u:/ is perceptually similar to their SHG /u:/ as they represent the same phonemic category in either NAE and SHG: they are both high back rounded vowels. Based on this, NAE /u:/ and SHG /u:/ are *similar* vowels, as they occupy the same phonological space, meaning there are no other phonemes found between the two in either language. Because their L1 vowel /u:/ competes with the foreign English /u:/, but the difference is so subtle, it will be less likely for a German native speaker to hear the acoustic difference between the two, which, in turn, can make its production difficult, due to perceptual similarity of these two sounds (see section on Speech Learning Model).

Two other NAE vowels, /ʌ/ and /ɒ/, will appear in the results section later, as these two proved themselves to be of interest during data analysis, even though they had not been previously reported as noteworthy in existing literature.

2.5.3 Voicing contrast of obstruents appearing in coda position

Another profound difference between the NAE and SHG phonological system is that SHG lacks a voicing contrast of obstruents in coda position. This means that in syllable-final position of words the SHG speaker produces voiceless obstruents only, even if the underlying coda obstruent is voiced, as reflected in the orthography:

Underlying voiced obstruents in coda position:

/b, d, g, v, z*/ → devoiced: /p, t, k, f, s/

SHG words		phonetic transcriptions
<i>stop</i>	<i>Laub</i>	[ʃt̥ɔp] [laʊp]
<i>Bart</i>	<i>Bad</i>	[ba:t] [ba:t]
<i>Sack</i>	<i>Tag</i>	[za:k] [ta:k]
<i>Schiff</i>	<i>oliv</i>	[ʃif] [o:lif]
<i>Maus**³</i>		[maʊs]

In some cases, final devoicing leads to two lexical items being indistinguishable, as the SHG examples of *Rat* ‘advice’ and *Rad* ‘wheel’ demonstrate. Consequently, when spoken in isolation, they are not reliably distinguished by native speakers/listeners of German (Brockhaus, 1995, p.3). In NAE, on the other hand, these obstruents in coda position are contrastive:

Voiceless obstruents in coda position /p, t, k, f, s, θ, ʃ, tʃ/

Voiced obstruents in coda position: /b, d, g, v, z, ð, ʒ, dʒ/

NAE words		phonetic transcription
<i>tripe</i>	<i>tribe</i>	[tɹaɪp] [tɹaɪ:b]
<i>fate</i>	<i>fade</i>	[feɪt] [feɪ:d]
<i>dock</i>	<i>dog</i>	[dɔk] [dɔ:g]
<i>chief</i>	<i>achieve</i>	[tʃi:f] [ə.tʃi:v]
<i>bus</i>	<i>buzz</i>	[bʌs] [bʌ:z]
<i>scythe</i>	<i>scythe</i>	[saɪθ] [saɪ:ð]
<i>rash</i>	<i>garage</i>	[ɹæʃ] [gə.ɹɜ:ʒ]
<i>letch</i>	<i>ledge</i>	[letʃ] [le:dʒ]

³ **SHG orthographic <s> is usually voiced in onset-position and syllable-medially /z/ in *sagen* [za:gn].

Therefore, the voicing contrast is a crucial cue, and while there are studies that support that NAE speakers rely predominantly on vowel duration of the vowel preceding the syllable-final obstruent – increasing vowel duration when the coda is voiced (Boersma, 2010) – it is generally reported that NAE distinguishes between voiceless and voiced coda obstruents (Reed & Levis, 2015 ; Hansen Edwards & Zampini, 2008 ; Archibald & Libben, 1995).

Previous studies have shown that native speakers of SHG may rely on the vowel duration differences of vowels preceding voiced and voiceless codas, respectively, but it is not a salient cue, nor do they actually produce a voicing contrast of the obstruent (Smith, Hayes-Harb, Bruss, & Harker, 2009 ; Roettger, Winter, Grawunder, Kerby & Grice, 2014). As these voiced and voiceless obstruent counterparts do not contrast in SHG, one can assume that this phonological process will be carried over from German, and NAE voiced coda obstruents will be devoiced by SHG L2 speakers, as has indeed been shown in several past studies (Smith et al., 2009).

2.5.4 Aspiration of voiceless stops in coda position in German

“Aspiration is a burst of air that follows the German plosives /p, t, k/, most often when they are in word-initial position” (Grantham O’Brien & Fagan, 2016). Likewise, NAE also aspirates stops in word-initial position (Gast & König, 2009). Aspiration on stops is a burst of air that follows the stop release. When aspiration is present, there is an increased voice onset time (VOT) between the stop release and the onset of voicing in the following vowel. This results in a sudden release of air (Grantham O’Brien & Fagan, 2016).

Aspirated [p^h, t^h, k^h] are allophonic variants of /p, t, k/ in German and English, the other allophones being [p, t, k]. In English and German, the aspiration takes place in initial position, but not when the voiceless stop is the non-initial member of a cluster, as demonstrated in the example below:

NAE: <i>pat</i>	[p ^h æt]	SHG: <i>Pakt</i> ‘deal’	[p ^h akt ^h]
NAE: <i>sport</i>	[spɔ:t]	SHG: <i>Sport</i>	[ʃpɔ:t ^h]

Aspiration additionally occurs in word-final position in German, where in English voiceless stops are usually not aspirated, and some of them are completely unreleased in this environment (König & Gast, 2009).

Aspiration in coda position would occur in, e.g. the following examples from German:

SHG:	<i>Trip</i>	trip	[tʁɪp ^h]
	<i>matt</i>	tired	[mʌt ^h]
	<i>Sack</i>	bag	[zʌk ^h]

In word-final positions, we find “the puff of air” of the final stop to be much stronger in German than in English (Bastug, n.d.). Consequently, German ESL learners will transfer the aspiration of /p,t,k/ sounds to English as well, resulting in productions like:

NAE:	<i>map</i>	SHG output:	[mæp ^h]
	<i>cat</i>		[kæt ^h]
	<i>lick</i>		[lɪk ^h]

[p^h, t^h, k^h] and [p, t, k] are allophones in SHG and the phonetic difference between these two sets of stops is minimal. According to the SLM, even though NAE and SHG have both aspirated and unaspirated [p,t,k] variants, they differ in coda position, where SHG exclusively has aspirated stops [p^h, t^h, k^h] as opposed to NAE. Therefore, in

coda position, these sounds can be considered *similar*, as the only difference is aspiration in coda-position in SHG, which NAE does not have. Therefore, the SLM predicts that the acquisition of unaspirated stops in coda position will be *difficult* for native speakers of SHG.

2.5.5 Substitution of NAE Interdentals /θ/ and /ð/ with SHG /s/ and /z/

/θ/ and /ð/ are interdental fricatives that are not present in the phonological inventories of most languages spoken around the world (Maddieson, 2013). They are thus cross-linguistically rare, or *marked*, and specific to very few languages. Due to their absence from most phonetic inventories, most L2 learners of English across the globe experience difficulties with their production (Rogerson-Revell, 2011). Their realization pattern can vary depending on the L1 of the ESL learner, but their substitution is consistent across languages (Lombardi 2003). For the native speaker of SHG, /s/ and /z/ substitutions are the most common, followed by /t/ and /d/ substitutions, respectively (L1 Dutch and Russian speakers favour /t/ and /d/ for NAE interdentals as well, for example) (Lombardi, 2003). In contrast, among native speakers of English, especially in children who have not yet acquired their interdentals, the most common substitutions are /f/ and /v/ (Yildiz, 2010). According to the Speech Learning Model, the phoneme /θ/ and its voiced counterpart /ð/ represent *new* sounds to the native speaker of German, as the SLM predicts that equivalence to any SHG sound cannot be established for the NAE interdentals (Honeybone & Salmons, 2015). According to previous studies, more inexperienced German native speakers will substitute the interdentals /θ/ and /ð/ with /s/ and /z/, respectively (Hancin-Bhatt, 1994 ; Hanulikova & Weber, 2012). Articulatorily, SHG /s/ and /z/ (alveolar) are the anatomically closest sounds to /θ/ and /ð/ (they are both

fricatives, the tongue is connecting with the teeth); they only differ in two aspects: the place of articulation and stridency. Both /s/ and /θ/, and their voiced counterparts /z/ and /ð/, are fricatives, so the manner of articulation is the same, but to articulate /s/ and /z/, the tip of the tongue is on the alveolar ridge, whereas for /θ/ and /ð/, it is placed between the upper and lower front teeth. For German native speakers, the acquisition of the English interdentalals can be “extremely challenging”, despite their classification as *new* sounds according to the SLM because the articulatory mechanisms involved in their production can be difficult to master (Hanulikova & Weber, 2012). Thus, a native speaker of German often produces an output like this:

NAE word	SHG speaker output:
<i>think</i>	[sɪŋk]
<i>although</i>	[ɒl'zəʊ]
<i>both</i>	[bəʊs]

2.5.1 Substitution of dark [ɫ] with light [l]

According to Baker and Goldstein (2008), “the use of only clear [l] or dark [ɫ], the substitution of one /l/ sound in place of the other, or the use of a non-English /l/ will not cause misunderstandings, but may contribute to non-native accent” (p. 116). The reason it will not cause misunderstandings is because dark [ɫ] is an allophonic variant of the light /l/ in the English language, meaning that they do not contrast phonemically. For comparison, in British English, either allophone occurs in specific positions. Light [l] usually precedes a vowel, which is why it is often found in syllable-onsets, such as *light* [laɪt], whereas [ɫ] appears after vowels, typically in syllable codas like in the word *milk* [mɪɫk]. When forming a clear [l], the tongue moves upward and forward towards the

alveolar ridge, whereas for dark [ɫ], the tongue moves upwards but backwards touching the soft palate (Reed & Levis, 2015).

However, in NAE, dark [ɫ] appears in all positions irrespective of syllable context and has a more vowel-like quality to it, which is why it is often vocalized in words like *school* [sku:ɔ] or *couple* [kʌpɫ], whereas an alveolar light [l] is not vocalized (Jones, 1966 ; Wells, 1982 ; Whitley, 2002 ; Barlow, 2014). German lacks this velar gesture of [ɫ] in all positions, suggesting that native speakers of SHG will substitute it with light [l] (König & Gast, 2009). With respect to predictions of the Speech Learning Model, this [ɫ] sound will be perceived as *similar* to speakers of SHG, and it will likely be ‘attracted’ by their native phonetic category of SHG /l/. Note that substitution of SHG [l] for NAE [ɫ] affect learners’ abilities to make lexical contrasts, because they do not differ phonemically, and some varieties of English do have [l] and [ɫ] phonetically, but they represent the same phonemic category. Thus, whether one says NAE [mɪɫk] or [mɪlk] does not affect the meaning.

2.5.2 Substitution of NAE /w/ with SHG /v/

Another very common pronunciation error that occurs in the speech production of native speakers of German learning English is the substitution of the NAE labiovelar approximant /w/ with the voiced labiodental fricative /v/ (Swan & Smith, 2001; Chamson, 2016). The German phonetic inventory lacks the sound /w/ but does have the sound /v/. Chamson (2016) predicts that less experienced German learners of English will consistently substitute NAE /w/ with SHG /v/, whereas in more experienced learners a different phenomenon could be observed: the more experienced group has phonetically acquired NAE /w/ as a *new* sound but overgeneralizes it and applies it to English /v/ and

/w/ words equally. This overgeneralization is typical in L2 speech. It occurs when a salient (noticeable) feature of a foreign language is detected by the ESL learner and perceived as ‘very English-sounding’, so they are under the impression that with frequent occurrence of the sound /w/ in their L2 speech production, they sound more authentic (Chamson 2016). Therefore, we may find patterns of underuse as well as overuse of [w] in the speech of native speakers of SHG:

underuse of [w]:

NAE: <i>wonder</i>	[wɒndə]	SHG output: *[vɒndə]
<i>west</i>	[wɛst]	*[vɛst]

overuse of [w]:

<i>village</i>	[vɪlɪdʒ]	*[wɪlɪdʒ]
<i>vain</i>	[veɪn]	*[weɪn]

Depending on proficiency and experience, some SHG speakers may substitute the /w/ in the cardinal direction ‘west’ with /v/ [vɛst], therefore turning it into a piece of clothing, namely a ‘vest’, whereas other SHG speakers will produce the piece of clothing ‘vest’ as [wɛst], therefore as the cardinal direction ‘west’. In either substitution case, /w/ and /v/ collapse to a single phonemic category where they lack contrast to the native speaker of German. Even though /w/ would represent a *new* sound according to the Speech Learning Model (Flege, 1995) that is unlike any other SHG sound, it shares some striking phonological similarities with the SHG sound /v/ (voicing, labial articulation, continuancy). However, /v/ lacks lip rounding and a dorsal gesture and is not approximant like /w/ is and can therefore not be considered very *similar*. Therefore, it should be easy to acquire for the SHG native speaker as a *new* sound.

2.5.3 NAE /ɹ/ and SHG /ʁ/

Generally, the most noticeable difference between NAE /ɹ/ and German /ʁ/ are the different places and manners of articulation. NAE /ɹ/ is an alveolar approximant (a liquid), whereas SHG /ʁ/ is a voiced uvular fricative (Grantham O'Brien & Fagan, 2016 ; König & Gast, 2009). In NAE, /ɹ/ represents a single phoneme, in SHG /ʁ/ represents the underlying phoneme which has a vocalized allophone, realized as dark schwa [ɐ] which occurs in postvocalic position (Grantham O'Brien & Fagan, 2016). To illustrate this process, one can refer to the example of British English, as BE, additionally to the approximant /ɹ/, has a vocalized allophone as well, namely [ə] which follows the same principle as that of SHG. The examples below demonstrate this:

SHG	<i>Bruder</i>	[bʁuːdɐ]	vocalized post-vocalically
NAE	<i>brother</i>	[bɹʌðɚ]	rhotacized post-vocalically
BE	<i>brother</i>	[bɹʌðə]	vocalized post-vocalically

(note that in German there are other allophonic variants of R present, such as alveolar trills /r/ or uvular trills /ʀ/, but for SHG, the uvular fricative is the most common one prevocalically)

Because NAE /ɹ/ and the underlying SHG /ʁ/ are articulated in different place and manner each, this sound of NAE can be classified as *new* for German learners of English and they generally perceive it as different from their own variant (Wode, 1980). As NAE /ɹ/ is unlike any other German sound, this is usually not confused with any other SHG phoneme (Wode, 1980 in James, 1988). Simultaneously, as Wode (1981) found, however, NAE /ɹ/ is acquired through different stages by SHG learners. He says that before the SHG speaker is able to produce NAE /ɹ/, he or she may go through different stages of approximating the SHG /ʁ/ to NAE /ɹ/:

- Stage 1: substitution with /ʁ/
 - Stage 2: substitution with /w/
 - Stage 3: substitution with a 'central frictionless continuant' /ɹ/
 - Stage 4: target-like retroflex or postalveolar /ɹ/
- (James, 1988)

If these substitutions occur, they will likely affect NAE /ɪ/ in prevocalic position. However, as SHG has vocalization of postvocalic /ʁ/, for postvocalic NAE /ɪ/, which is not vocalized, the speaker may apply this pattern of the L1 to his or her L2 as well, as one study revealed (Ulbrich & Ordin, 2014). Thus, in postvocalic position, it can be predicted that the native speaker of German may vocalize the NAE /ɪ/ instead of substituting it with one of the approximations Wode (1980) mentions. Even though the different stages of acquisition illustrated above suggest that NAE /ɪ/ will initially be difficult to articulate, NAE /ɪ/ as an alveolar approximant is unlike any other SHG sound, with its closest counterpart being the rhotic uvular fricative /ʁ/, so phonetically, /ɪ/ would represent a *new* sound to the native speaker of SHG. Consequently, the SLM predicts that /ɪ/ is *easy* and would eventually be acquired by the native speaker of German.

In this section on NAE /ɪ/ it has been described that two different substitution phenomena can occur in the speech of native speakers of SHG. Prevocalic /ɪ/ differs from that of the German uvular fricative /ʁ/ and therefore represents a new sound to be acquired by the native speaker of SHG. Postvocalic NAE /ɪ/, however may be treated differently, as German only has vocalized [ɐ] in that position. This may then be transferred to English.

2.5.4 Speech Learning Model (Part II): Predictions for NAE Segments

On the basis of comparing the sound systems of NAE and SHG and in regards to Flege's SLM, one can establish the predicted difficulty for the NAE sounds that the German learner would face. As new sounds are predicted to be easy to acquire and similar sounds are predicted to be difficult, one first has to establish the criteria that classify sounds as new or similar. As mentioned at the start of this section, I adopted the hypothesis that

sounds that differed in one phonetic property were *similar*, while those that differed in more than one were *new*. Of the sounds discussed in Section 2.5, three were classified as *similar*: NAE /u:/ and SHG /u:/ (front/backness), NAE [ɥ] and /l/ (velar constriction), NAE [p̄, t̄, k̄] and SHG [p^h, t^h, k^h] (release of stop closure). On the other hand five were classified as *new*: NAE /æ/ and SHG /ɛ/ (vowel height (F1) and frontness (F2)), NAE interdentals /θ/ - /ð/ and SHG /s/ - /z/ (interdental vs. alveolar ; acoustic intensity [stridence]), NAE /w/ and SHG /v/ (glide vs. fricative ; labiovelar vs. labiodental) , NAE /ɹ/ and SHG /ʁ/ (liquid vs. fricative ; alveolar vs. uvular).

Summarizing the predictions of the SLM on difficulty, the following chart comparing NAE to SHG sound segments in terms of *new* and *similar* was generated:

Table 2.3 SLM Classifications of NAE and SHG sounds⁴

NAE sound	/æ/	/u:/	[p̄, t̄, k̄]	/θ/ - /ð/	[ɥ]	/w/	/ɹ/
closest SHG sound	/ɛ/	/u:/	[p ^h , t ^h , k ^h]	/s/ - /z/	/l/	/v/	/ʁ/
phonetic similarity	new	similar	similar	new	similar	new	new
SLM prediction	easy	difficult	difficult	easy	difficult	easy	easy

⁴ Note that NAE voicing contrasts of obstruents was not considered for prediction of difficulty in the section on the SLM, as this is a **phonological** process. SHG has /b,d,g,z,v/ in onset and syllable-medially, and the SLM concerns itself with the formation of *new* and *similar* L2 **phonetic** categories. German does have a phonetic category for these sounds, so contrasting them phonetically is not a problem. However, contrasting them phonologically would be, as SHG devoices all underlying voiced coda obstruents.

Coda aspiration is a phonological process in German that NAE does not have. The NAE unreleased stops [p̄, t̄, k̄] would be *similar* to the native speaker of SHG, as they are still allophones of the same stops, although they are phonetically absent from SHG altogether.

After analysis of the results the NAE speech sounds for accuracy in M.S.'s production across three tasks, the predictions of the SLM will be discussed in Chapter 5 to demonstrate whether phonetic similarity or dissimilarity can account for difficulties in acquiring these NAE sounds in M.S.'s case.

Chapter 3 Methodology

3.1 Participants

This project involved two participants: A native speaker of German, and a native speaker of North American English.

The native German speaker (henceforth “M.S.”) was the experimental participant. He is 50 years old, male, and has lived in North America for 25 years. He resides in Victoria, British Columbia, Canada. He speaks English with a distinct German accent.

The native English speaker (henceforth “C.E.”) was the control participant. He is 52 years old, male, and currently lives in Victoria, B.C. This participant was important for two reasons: he served as a comparison to M.S. for the native speaker judges who served to assess M.S.’s accentedness, intelligibility, and comprehensibility. Secondly, he was the native speaker baseline for some of the segmental productions that had not been reported in the literature review and emerged from the production data, as well as for segmental productions where a direct comparison with a native speaker was needed to assess the differences between M.S.’s productions and L1 production.

3.2 Instruments and Stimuli

To obtain a picture of M.S.’s production of NAE segments, 3 different experimental tasks had to be designed and administered to M.S. and the control participant C.E. These three tasks were 1) a wordlist (most formal), 2) a sentence reading task, and 3) a semi-spontaneous production task (least formal). C.E. only participated in the wordlist and sentence production task, as the semi-spontaneous task consisted of a set of questions that M.S. responded to, so his answers could not be controlled for. Additionally, M.S. filled

out a questionnaire on his language background and experience with English and partook in a 30-minute interview that included a self-assessment of his pronunciation.

Questionnaire

Before M.S.'s production data was collected, he filled out a **questionnaire** on his personal, educational, and language background to get a description of his language learning experience (Lin, 2013) (**APPENDIX A**), to assess the quality of the input of English he had received and what could be contributing factors to his pronunciation proficiency, as knowledge of other languages and the age of acquisition as well as length in English speaking countries.

Interview Self-Assessment

Additionally, a 30-minute oral interview served as a means to obtain further information from M.S. about his pronunciation of NAE which was focused on the acquisition of specific sounds and his experience living in an English-speaking country. The interview contained a component of self-assessment in which M.S. commented on his pronunciation of English sounds to bring to the surface potential strengths and weaknesses of his pronunciation and to find out whether these reported strengths and weaknesses correlate with the findings reported in the results section. Derwing (2003) conducted a study in which immigrants to Canada were interviewed on their perceived pronunciation ability and found that the subjects were able to provide comments on pronunciation difficulties pertaining to "salient segmental units". Therefore, the interview contained questions on how M.S. assessed his own pronunciation abilities in English in regards to specific sounds.

This additional information from the subject himself would be added to the analysis of the production data. This element further served to gain insight into extra-linguistic factors that may contribute to his learning of NAE sounds and to his perception of being a foreign language learner in a mainly English-speaking Canadian culture with his roots and attachment to his home country Germany. The goal was to learn about M.S.'s own motivation and attitude towards learning a foreign language and how these factors influence his pronunciation. Achirri (2017) utilized an interview and a questionnaire to obtain data from the subject on their foreign accentedness in relation to their identity, reporting that her subject never attempted accent-reduction because she used her accent an identity marker. The interview brought to the surface that the ESL learner was able to make assessments on the fossilization of their pronunciation and that their pronunciation would no longer change in some areas (Achirri, 2017). Thus, these interviews and questionnaires served to elicit information on the subject's pronunciation background, views, and identity as a learner addressing problematic areas of pronunciation and why accentedness matters – or does not matter.

These approaches of self-assessment from interviews were utilized in this study.

Stimuli

Wordlist Task

The wordlist reading task consisted of a wordlist of 300 items (tokens) that were high frequency words in English. The tokens were high-frequency words to avoid exacerbating lexical access issues that would lead to production errors due to unfamiliarity rather than articulatory difficulty. These tokens contained a number of segments that have been identified in the literature to be problematic for speakers of SHG

(detailed in Section 2.5, and for more examples refer to **APPENDIX B**). Some of the high frequency words contained one critical segment, such as /v/ in *evidence*, whereas others contained multiple segments of interest. For example, the stimulus NAE *collateral* contains four segments that could potentially pose difficulties to M.S., namely two dark [ɫ]s, NAE /ɪ/, and the NAE vowel /æ/. The researcher tried to ensure that, if a segment was known to be problematic, sufficient words containing it were presented to the subject, although this was difficult to control for each segment. As a number of potentially problematic segments were targeted, and to avoid an even bigger word list, the problematic segments were not controlled in terms of equal appearances in a specific phonetic or syllabic environment (e.g., a segment may appear 45 times in onset, but 56 times in coda).

Sentence Reading Task

The sentence reading task was comprised of 60 sentences with a length of 5-12 words each. Those sentences were generated by the researcher and contained most of the words from the wordlist. Both, M.S. and C.E., were asked to read out sentences containing the critical segments. Like the wordlist words, some sentences contained a number of segments under focus, e.g. “Leaning over the ledge reverses Batman’s naïve rearview mirror” (**APPENDIX C**). There was some overlap between the items that appeared in the wordlist and the sentences, thus some words (roughly 220) of the wordlist items made it into the sentences as well. The overlap between wordlist and sentences meant that a direct comparison could be made to look at how task type could affect performance.

Semi-Spontaneous Production Task

The least formal production task that was generated for the purpose of obtaining natural speech production data from M.S. was a short interview with the researcher that lasted ~10 minutes and which served to elicit spontaneous speech production without any prior input, which would then be the most reflective of his L2 pronunciation in a real life situation. Among the interview questions were items such as “What aspects of Canadian culture appeal to you the most?” or “What is your most memorable experience coming to Canada?”.

Native Speaker Judgments (NSJs)

In addition to the production data and the elements used to solicit information from M.S. about himself and his language background and experience, two native speaker judges (henceforth: NSJs) were asked to assess M.S.’s pronunciation ability on a global level and to provide goodness ratings on his speech production that would pertain to the key concepts of intelligibility, comprehensibility, and accentedness. In order to do so, the judges listened to a part of the production data (the same 9 utterances from the reading task for both NSJs, ranging from 4-10 words per utterance) produced by both M.S. and C.E. and were asked to orthographically transcribe the heard utterance on a spreadsheet. Immediately afterwards they listened to the utterance again and provided a rating for comprehensibility with a score on a scale from 1-9 of how easy they found the utterance to understand (1=extremely easy to understand; 9 = extremely difficult to understand). In a separate session a week later, they listened to the 9 utterances again and provided ratings on the perceived accentedness of the productions they heard. The judges once again had a sheet with spaces for the utterances without orthographic input and were

asked to provide a rating on foreign accentedness on another 9-Likert scale (1= no foreign accent; 9 = very thick foreign accent). Getting accentedness ratings for individual utterances served to establish whether accented production correlates with intelligibility and comprehensibility of M.S.'s productions and whether segmental errors would contribute to perceived foreign accented or impoverished communication skills.

Additionally, the NSJs provided qualitative comments on certain segmental productions obtained through a questionnaire (designed after Murphy, 2014) (**APPENDIX E**). The sentences M.S. and C.E. read out were relatively free of semantic priming, meaning that arguments of the sentence, or adjacent words, did not telegraph their meaning through context. This also meant that the judges listened to connected speech that did not allow them anticipate the context of each word that was presented to them: e.g., “Chewing polar bear shaped gum for partners should be key”, so as to avoid the judges inferring a word rather than perceiving one.

3.3 Procedure

Upon receiving ethics approval from the University of Victoria, the experimental and control subjects were recruited to participate in the study. In my capacity as a researcher I approached potential subjects who had previously expressed interest in participating in a language experiment. Especially M.S. was enthusiastic about learning something about his accent and pronunciation, as he pointed out to me, so he was asked to join for 2 individual sessions to obtain the production data, the questionnaire and the interview (self-assessment).

M.S. completed the background questionnaire on his language experience and exposure to English before the recording took place. The recording took place in a sound-attenuated booth at the Speech Research Lab of the Department of Linguistics at the University of Victoria using a M-Audio Luna microphone to record digital sound files onto a Mac desktop computer in WAV format at a 44.1 KHz sample rate of 16 bits per sample. M.S. read the 60 sentences that were presented to him one at a time on a Power Point Presentation. At the end of each sentence, the researcher would manually change the slide so he was able to read them at a pace that he found comfortable. After completing the sentence task reading, which took ~10 minutes, M.S. was able to take a short break, after which he completed the wordlist task. Each item was presented to him one at a time on a Power point slide that was manually controlled and turned to the next slide after each word production. The wordlist reading task took ~12 minutes for all 300 tokens.

The control C.E. joined me at the lab on a different day and completed the same tasks in the same order. The time dedicated to the sentence production and wordlist task was slightly less than that of M.S., as C.E.'s speech rate was significantly faster for the sentences than that of M.S.: ~6 minutes for the sentence task, and ~10 minutes for the wordlist task.

After completing the formal task types, M.S. joined me at the Speech Lab for another recording session which was comprised of two parts of an interview. The first part of the interview was the semi-spontaneous speech production task, which took roughly 10 minutes and was recorded using the same equipment. The second part of the interview was not used for production analysis but served as a self-assessment to evaluate

M.S.'s English pronunciation and speak about extra-linguistic factors (**APPENDIX D**) that may contribute to his production. This part took ~15 minutes. With the completion of this interview, M.S.'s involvement in the study was complete.

After obtaining the production data from M.S. and C.E., the native speaker judges were invited to participate in the listening experiment, for which they provided accentedness and comprehensibility ratings assigned to each speaker (intelligibility was calculated during data analysis). The NSJs rated on two different occasions a week apart. The first time they met with me at the Speech Research Lab at the Linguistics Department at the University of Victoria and agreed to dedicate 20 minutes of their time to listen to the 9 utterances by M.S. and C.E., respectively. Both judges received pair of headphones (Sony-MDR-V150) and listened to the production data and orthographically transcribed each utterance on a spreadsheet. After each utterance, the researcher paused the recording, so they had sufficient time to write down what they heard. Each utterance could be listened to multiple times, if necessary. Underneath the utterances they had transcribed was a Likert scale of 1-9, and they were asked to circle the number that they felt would most accurately describe the effort they had to put to understand the utterance. A week later, they were asked to join me at the lab again and they listened to the same 9 utterances to provide accentedness ratings on a scale from 1-9 (explained above). This took roughly 10 minutes of their time.

After the completion of all production experiments and the native speaker judgment sessions, a trained phonetician and myself transcribed all production data from the wordlist, sentences, and M.S.'s interview. Disagreements in transcriptions between

the researcher and the transcriber were discussed and a common transcription decided upon.

3.4 Data Analysis

Analysis of M.S. and C.E.'s productions involved two components: an auditory analysis (transcription) done by the researcher with the aid of a native English-speaking trained phonetician, and an acoustic analysis. All of C.E. and M.S.'s words from the wordlist and sentences were transcribed, as was each word of M.S.'s interview. The transcriptions were narrow transcriptions. Data about specific productions of words, and especially of segments, were taken from these narrow transcriptions, and if needed compared directly to the sound files, and visual displays of the acoustic signal: waveform and spectrogram. Data was collected about each segment of interest and tabulated.

Acoustic analysis was done for segments for which auditory analysis alone was not sufficient to determine all the details of production variance. This meant mainly vowels, which vary on a continuum, as opposed to consonants, which are mainly categorical.

Acoustic analysis of vowels usually measures quality, and this relies mainly on properties known as formants. The first two of these, called F1 and F2, are the formants associated most unambiguously with the specific articulatory components of vowels (Kent and Read, 2002). F1 is inversely associated with tongue/jaw height (e.g., for the sound [i] the tongue is near the roof of the mouth and F1 is around 300 ~ 400 Hz, but for [a] the tongue is at the bottom of the mouth and the jaw is open, and F1 is around 700~800 Hz). F2 is associated with tongue frontness (e.g., for the vowel [i] the tongue is

advanced in the mouth and F2 is around 1900~2000 Hz, but with [u] the tongue body is retracted in the mouth and F2 around 1100~1000 Hz). Measurements of these formants allow us to make inferences about where in the mouth a given vowel is articulated and what the differences are across vowels (Kent & Read, 2002 ; Ladefoged, 2006).

Consonants usually (at least for this study) do not vary on a continuum like vowels do, and for this reason I did not do acoustic analysis. Consonants are also diverse in their acoustic properties and this makes it difficult to compare acoustic measurements across different types of consonants. It also means that the differences are most often sufficiently analysed auditorily (Kent & Read, 2002). For example, there is nothing contrastive falling between [p] and [t] in NAE or SHG, so the acoustic differences between the two do not need to be investigated the same way the differences between [æ] and [ɛ] need to be. Therefore, a phonetician could transcribe fine differences in consonants, whereas for vowels acoustic measurements would be needed to confidently report fine phonetic detail. Consequently, this dual way of measuring vowels and consonants was adopted to allow for optimal analysis.

Vowels were acoustically analyzed in Praat (Boersma & Weenink, 2018) and the first and second formants (F1 and F2) were measured in Hertz at the midpoint of the vowel, vowel boundaries being determined by the onset or offset of regular pitch pulses with multiple formants (at least two formants). Formant measurements were extracted from Praat into Microsoft Excel, and then imported into R (R Core Team, 2018), where they were plotted using ggplot2 (Wickham, 2009).

For any segment for which there appeared to be a trend, and for which strong confidence in the apparent trend was desired, a Welch's 2-sample t-test or a Pearson's

Chi-squared test were used (depending on the nature of the variables) to determine statistical significance. This was also used for testing between-NSJ factors with the Native Speaker Judges (see below). Statistical tests were not used for most segments, because for most of them clear trends were visible in the data and the numbers were robust enough that it was not felt to be warranted, or no obvious trend was visible and tests were likewise not felt to be warranted.

The Native Speaker Judgments were analyzed by comparing cross-judge ratings of comprehensibility and accentedness for each utterance as produced by M.S. and compared to those of C.E. For example, the accentedness rating given by NSJs to an utterance from M.S. was compared to the rating by the NSJs of the same utterance from C.E.; the same was done for comprehensibility scores. Intelligibility was determined by counting the number of words the judges incorrectly transcribed or left out in each utterance, e.g., the intelligibility rate of an utterance would be the number of errors in the transcription by the number of words total in the utterance. Statistical significance between the intelligibility, comprehensibility, and accentedness ratings of M.S. and C.E. (between speaker relation) were determined using a Welch's 2 Sample t-test and the correlation (if any) of these ratings to each other (e.g., accentedness with intelligibility; intra-speaker relation) with a Pearson's correlation.

Chapter 4 Results

In this chapter the results of the three production tasks that M.S. underwent (wordlist, sentences, and semi-spontaneous interview) will be reported in light of the differing intra-linguistic factors described in Chapter 2 (task type, orthography, cognate status, vowel environment, syllable context, and stress). The phonetician and the researcher (myself) transcribed the production data from the wordlist, sentences, and semi-spontaneous production task until consensus was reached. Once the accuracy scores adopting a threshold of >80% (Eckman, 1991) were calculated for each task, the scores were compared across task type to see if task formality had an effect on production accuracy. Following the calculation of accuracy scores, intra-linguistic factors that could have a bearing on M.S.'s production performance were examined (4.1 – 4.12). Only the factors that were considered to be relevant for each segment and played a clear role in the production of the target sounds, respectively, were reported. Finally, this chapter wraps up with a section on the results of the native speaker judgments on intelligibility, comprehensibility, and accentedness (4.13).

4.1 Results on Substitution of /æ/

As reported in Chapter 2, the NAE low front vowel /æ/ was reported to pose difficulties to native speakers of German, as German lacks this vowel that is phonetically situated between SHG /ɛ/ and SHG /a:/. In the three production tasks, a total of 219 tokens of NAE /æ/ were embedded into English high frequency words that M.S. produced to obtain insight into the acquisition of this *new* NAE vowel that is absent from SHG.

The table below shows the total of M.S.'s productions of target /æ/ for each task, respectively, listing his substitutions from the most favoured to the least in descending order.

Table 4.1 M.S.'s productions of /æ/

Target	Production	Wordlist (T1)	Sentence (T2)	Semi-spontaneous (T3)	Total
	<i>Total</i>	<i>51</i>	<i>70</i>	<i>98</i>	<i>219</i>
/æ/	[æ]	27 (53%)	20 (29%)	39 (40%)	86 (39%)
	[ɶ]	16 (31%)	11 (16%)	14 (14%)	41 (19%)
	[ɛ]	4 (8%)	31 (44%)	41 (42%)	76 (35%)
	[a]	1 (2%)	7 (10%)	4 (4%)	12 (5%)
	[ə]	2 (4%)			2 (1%)
	[e]	1 (2%)			1 (<1%)
	[ɪ]		1 (1%)		1 (<1%)

Task 1 Wordlist

In the wordlist task, M.S. produced 51 /æ/, of which 27 were on target, an accuracy of 53%. Substitutions included four [ɛ] (8%), 16 [ɶ] (31%), two [ə] (4%), and one each [a:] (2%) and [e] (2%).

Task 2 Sentences

In the sentence task, M.S. produced 70 instances of /æ/, of which 20 were on target, an accuracy score of 29%. Also produced were 31 tokens of [ɛ] (44%), 11 tokens of a raised [ɶ] (16%), seven of [a:] (10%), and one token of [ɪ] (1%). Note that four productions of target /æ/ that were produced as [ɛ] would be acceptable in varieties of North American English. They occurred in words where /æ/ was followed by [ŋ], as in *drank* or *angry*,

and in this environment /æ/ tenses, raises and is realized as [e~ɛ] in a large number of NAE varieties, including Canadian English (Boberg, 2015).

Task 3

In the semi-spontaneous production task M.S. produced 99 instances of /æ/, out of which 39 were on target, an accuracy rate/score of 40%. Also produced were 41 tokens of [ɛ] (41%), 14 of [æ̃] (14%), and four of [a:] (4%), and one instance of deletion (1%) in the word *program*.

Across all three tasks 219 instances of /æ/ were produced, out of which 86 were on target, an overall accuracy of 39%. The most common substitution type was [ɛ] in 76 cases (35%), meaning that the rate of [ɛ] productions was similar to that of [æ̃]. The next most common substitution was [æ̃] in 41 cases (19%), followed by the substitution [a:] in 12 cases (5%). There were also two instances of [ə], one instance of [e], and one instance of [ɪ], as well as one deletion.

The plot below shows the mean value for all of M.S.'s various /æ/ realizations (hollow square), and ellipses surrounding these realizations, and compares them to a mean value and ellipse of standard values for NAE [æ] (the smaller, more compact ellipse, values from Kent & Read, 2002⁵).

⁵ NAE vowel comparison measurements were taken from Kent and Read (2002), as they reported formant values for a number of North American English speakers; as M.S.'s productions are being compared to NAE, it was felt that they better served as an L1 comparison than the values from one Western Canadian English speaker (C.E.).

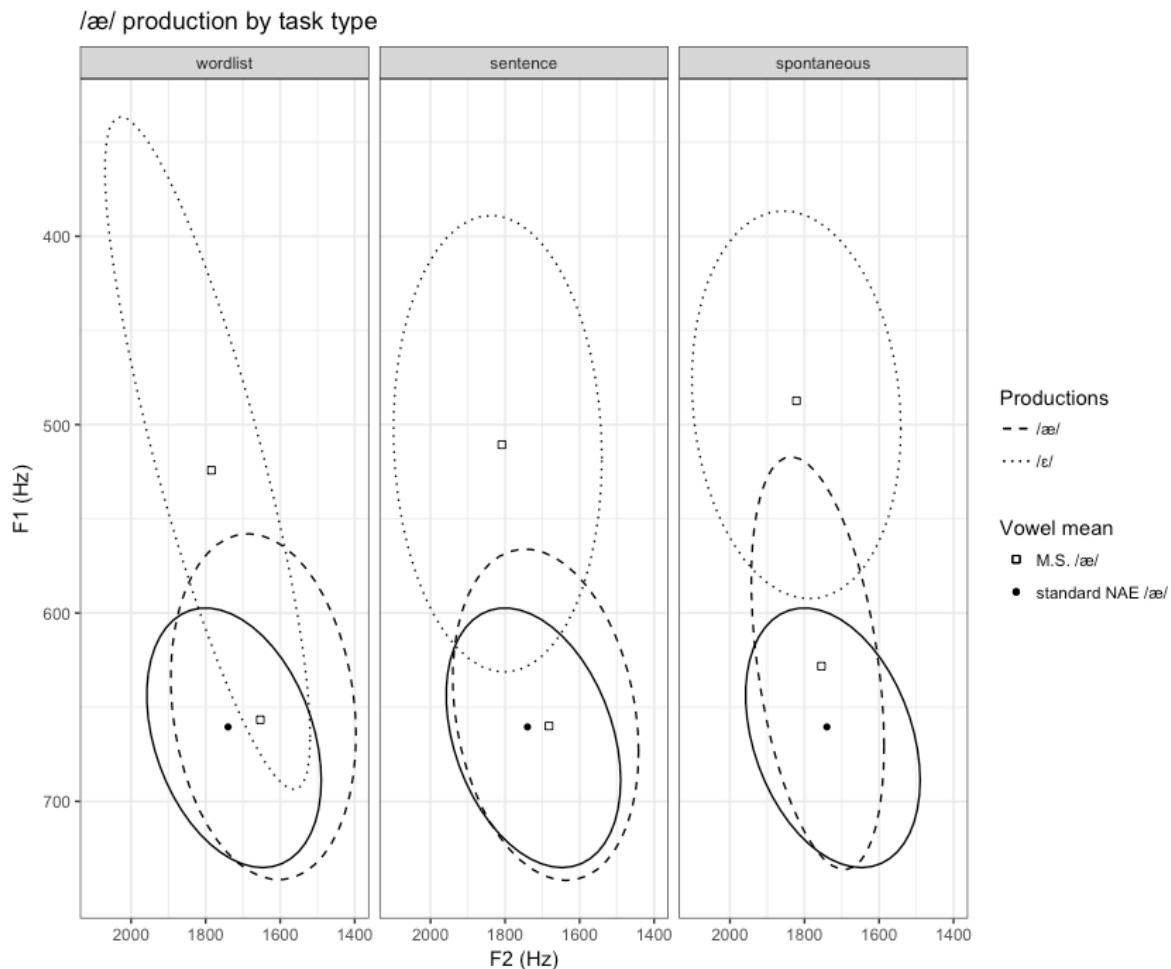


Figure 4.1 M.S.'s Acoustic productions of /æ/ compared to NAE standards

As task formality decreases, the mean F1 of [ɛ] (the dotted ellipse) slightly lowers, and the ellipse grows more compact (representing the greater number of productions in the Sentence and Semi-spontaneous tasks). Overall though, productions of [ɛ] for /æ/ are relatively acoustically consistent across tasks. For the wordlist and the sentence tasks, [æ] (the dashed ellipse) remains stable but begins to also have mean F1 lower in the semi-spontaneous task, as well as have more variable production (as seen with the distorting ellipse). It is not, however, too distant from the standard NAE [æ] values (the solid ellipse).

Task Type

Comparing accuracy scores across the three different task types, the wordlist task showed the highest accuracy at 53%, followed by the semi-spontaneous task with an accuracy of 40%, and the sentence task at 29%. Between the wordlist and the sentence tasks, 34 words containing /æ/ could be directly compared, as they appeared in both tasks. 10 words were on-target in both wordlist and sentences, 8 were off-target in both the wordlist and sentences, 3 were off-target in the wordlist but on-target in the sentences, and 13 were on-target in the wordlist but off-target in the sentences. This means that M.S. becomes less target-like as the task becomes less formal. This could also be due to the influence of other words in the sentences functioning as distractors forcing M.S. to focus his attention on other sounds than in the wordlist.

Cognate Status

Some of the English tokens M.S. produced have German cognates that are spelled similarly, e.g., NAE *accent* and SHG *Akzent* or NAE *salad* and SHG *Salat* where English has /æ/, German has /a:/. Therefore, if the cognate status influenced M.S.'s production of these tokens in English, one would suspect that he would have replaced [æ] in NAE *accent* and NAE *salad* with [a:]. However, this was not the case. When M.S. substituted [æ] with [a:], this seemed to be for different reasons. In order to account for these [a:] productions of target /æ/, a different phenomenon observed in varieties of British English should be taken into account: The Received Pronunciation (RP) variety of British English and a number of Australian varieties exhibit what one commonly refers to as *trap-bath split* (Bauer, 2015 ; Reed & Levis, 2015), but nearly all of North American English does not. For varieties of English with this split a number of lexical items that at one point

historically were produced with [æ] retracted to varying degrees of [a~ɑ] (Wells, 1982). The split marks the difference between orthographic <a> being pronounced as [a:] before the fricatives /f/, /s/, and /θ/ in varieties of British English, as in the words *shaft*, *last*, and *bath*, but as [æ] in all other instances, e.g.: *cat*, *trap*, *can*. However British varieties pronounce /æ/ as [a:] in specific environments: (*chance* / __ns/, *castle* / __sl/*advantage* / __nt/, *demand* / __nd/, *branch* / __n(t)f/, *example* / __mpl/ (Wells, 1982). Most varieties of NAE do not have this split, so <a> is pronounced as [æ] regardless of environment.

Interestingly enough, almost all tokens that would be affected by the *trap-bath split* in British English (BE), such as *after*, *grass*, or *last*, instead of [æ] as in NAE, M.S. produced with [a:]. This pattern was only found in the more formal tasks, but not in the informal task, except for one instance of *ask*. As M.S.'s instruction of English at a German high school followed the framework of British pronunciation, and since he had spent several years in Australia, this can possibly account for him acquiring the split through instruction and residence in a country in which a variety of English with the split is spoken (Wells, 1982).

4.2 Substitution of NAE /ʌ/ with SHG /a/

The low-mid back/central unrounded vowel /ʌ/ can be regarded as the stressed, and therefore “twin” variant of NAE /ə/. SHG also has /ə/ in unstressed syllables, but it lacks the stressed counterpart /ʌ/. Acoustically, /ə/ and /ʌ/ can be classified as *similar*, which is why the SLM would predict difficulty to acquire this sound phonetically for a native speaker of German. When listening and transcribing the data, M.S.'s speech

production seemed to exhibit a high amount of on-target productions of [ʌ], which was confirmed through acoustic analysis of the vowel in 109 instances that M.S. produced.

Low-mid back/central unrounded vowel /ʌ/ has two productions across tasks in M.S.'s speech: on-target [ʌ], like in [stʌdi], /stʌdi/ *study*, and substitution [a], like in [stadi], /stʌdi/ *study*. On-target productions dominated, accounting for about two-thirds of overall productions. These are laid out in the below table:

Table 4.2 M.S.'s productions of /ʌ/

Target	Production	Wordlist (T1)	Sentence (T2)	Semi-spontaneous (T3)	Total
/ʌ/	<i>Total</i>	24	57	28	109
	[ʌ]	20 (83%)	34 (60%)	17 (61%)	71 (65%)
	[a]	4 (17%)	23 (40%)	11 (39%)	38 (35%)

Tasks

M.S. produced 20 out of 24 productions of /ʌ/ as on-target in the wordlist condition, an accuracy of 83%. In the sentence task he produced 34 of 57 productions on-target, an accuracy of 60%. In the semi-spontaneous task he produced 17 out of 28 on-target [ʌ], an accuracy of 61%. In all tasks the only substitution was low central unrounded vowel [a].

The following figure presents the acoustic measurements of M.S.'s /ʌ/ compared to the standard NAE /ʌ/. The vowel spaces are outlined in ellipses, with the solid ellipse surrounding the standard productions, the dashed ellipse surrounding M.S.'s [ʌ], and the dotted ellipse M.S.'s [a]. Means are represented by a hollow square for the standards and filled circles for M.S.'s productions.

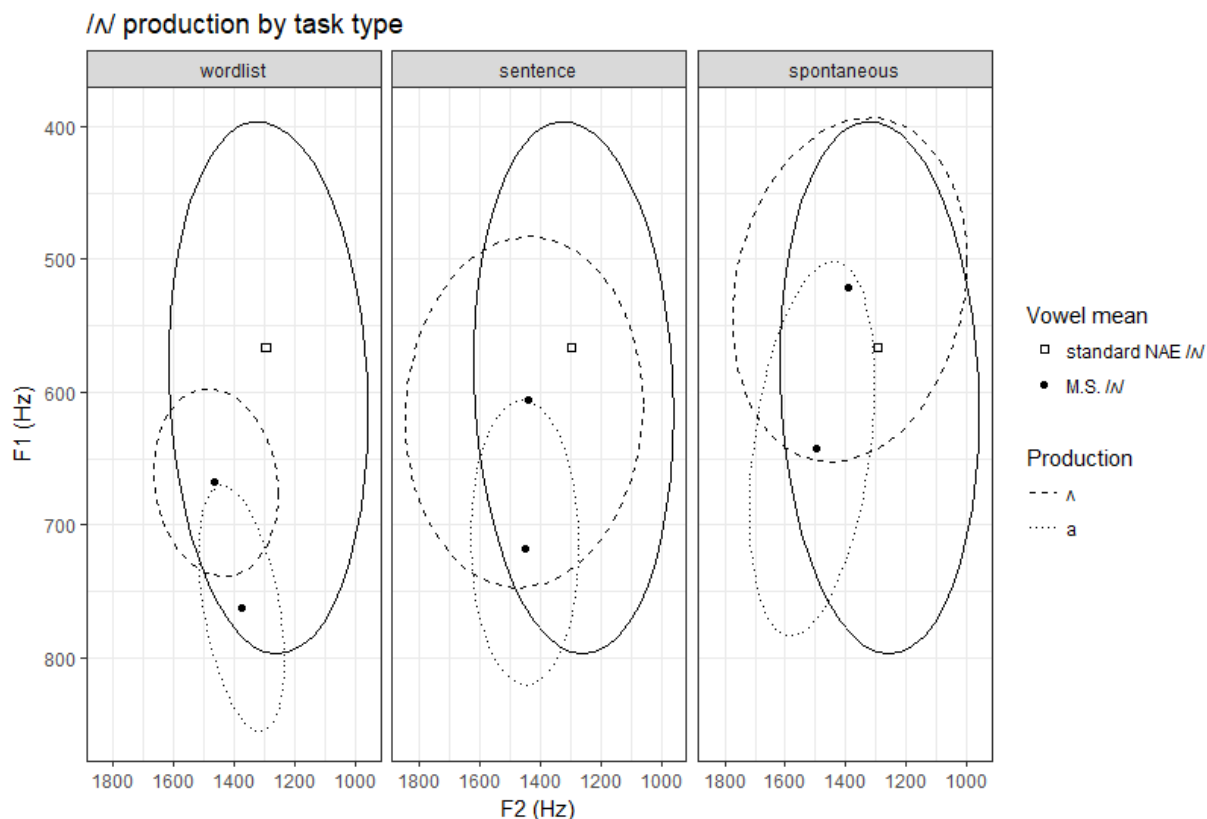


Figure 4.2 M.S.'s Acoustic productions of /ʌ/

Acoustically, all of M.S. productions were slightly fronted relative to the space of the standard productions of [ʌ] in NAE, with higher F2s. Additionally, the acoustic space of M.S.'s productions shifted “up” and right as task formality decreased, that is, average F1 and F2 decreased. This meant that productions transcribed as [ʌ] from the wordlist task were acoustically similar in terms of F1 to productions later transcribed as [a] in the semi-spontaneous task! Perhaps this is an effect of the relative nature of the two, with the productions transcribed as [a] being consistently lower (i.e., had higher F1s) than the [ʌ] productions, and the discrimination based on this relative distinction. The acoustic analysis suggests that M.S. is working to produce [ʌ] and is possibly using /a/ as the baseline, indicated by the production of tokens as [a]. /a/ being the closest SHG vowel to

/ʌ/, M.S. uses [a] as a basis and has centralized this vowel to produce [ʌ], accounting for the slight frontedness of his productions and the ambiguity of [ʌ]~[a].

4.3 On-target /ɒ/

This NAE low back vowel /ɒ/ was not originally reported to be problematic for native speakers of SHG according to existing literature. However, during data collection this sound became of interest, due to the fact that it was perceptually on target to the phonetician and the researcher, which necessitated acoustic analysis. Interestingly enough, SHG lacks this sound, its closest equivalent being SHG /ɔ/, which is also found in some varieties of English, but not typically in NAE, and specifically Canadian English. According to the SLM, this sound could be classified as a *similar* vowel.

The table below shows all 119 instances of NAE [ɒ] embedded into tokens of the production stimuli, their accuracy scores M.S. achieved for each task, and their substitutions:

Table 4.3 M.S.'s productions of /ɒ/

Target	Production	Wordlist (T1)	Sentence (T2)	Semi-spontaneous (T3)	Total
/ɒ/	<i>Total</i>	41	46	32	119
	[ɒ]	41 (100%)	40 (87%)	28 (88%)	109 (92%)
	[ʊ]		3 (6.5%)		3 (3%)
	[ʌ]		3 (6.5%)	4 (12%)	7 (5%)

Task 1

/ɒ/ was on-target 41 out of 41 times, that is, 100% of the time, in the wordlist task.

Task 2

M.S. was on-target 40 out of 46 times in the sentence task, 87%. He substituted [ʊ] and [ʌ] three times each.

Task 3

In the semi-spontaneous task M.S. was on-target 88% of the time, 27 out of 32 times. He substituted [ʌ] four times.

The productions of [ʊ] in the sentences, while unexpected and noticeably different, are explainable. They occurred in unstressed words both written with <o>: *on* and *not*. It is likely that the reduction of vowel quality associated with the lack of stress on these tokens accounts for the off target production. Ultimately, they are not indicative of a larger trend in M.S.'s /ɒ/ production, which is overall very on-target to the perception of the transcribers and can be considered “acquired”.

The data suggested that M.S.'s /ɒ/ is on-target, meaning that one cannot detect a difference perceptually between M.S.'s [ɒ] and a L1 English speaker's [ɒ], a finding that is borne out in the acoustic evidence. The slight decrease in accuracy across task is also evident in the acoustic data, as the mean /ɒ/ moves further away from the standard values (the stationary ellipse; values from Kent & Read, 2002) with decreasing task formality. The plot below illustrates this. As with /æ/, the three panels show ellipses containing the productions of /ɒ/ by M.S. across the three tasks, with a mean represented by a hollow square, and the standard productions of NAE /ɒ/ reported in Kent & Read (2002), with a mean represented by a solid circle.

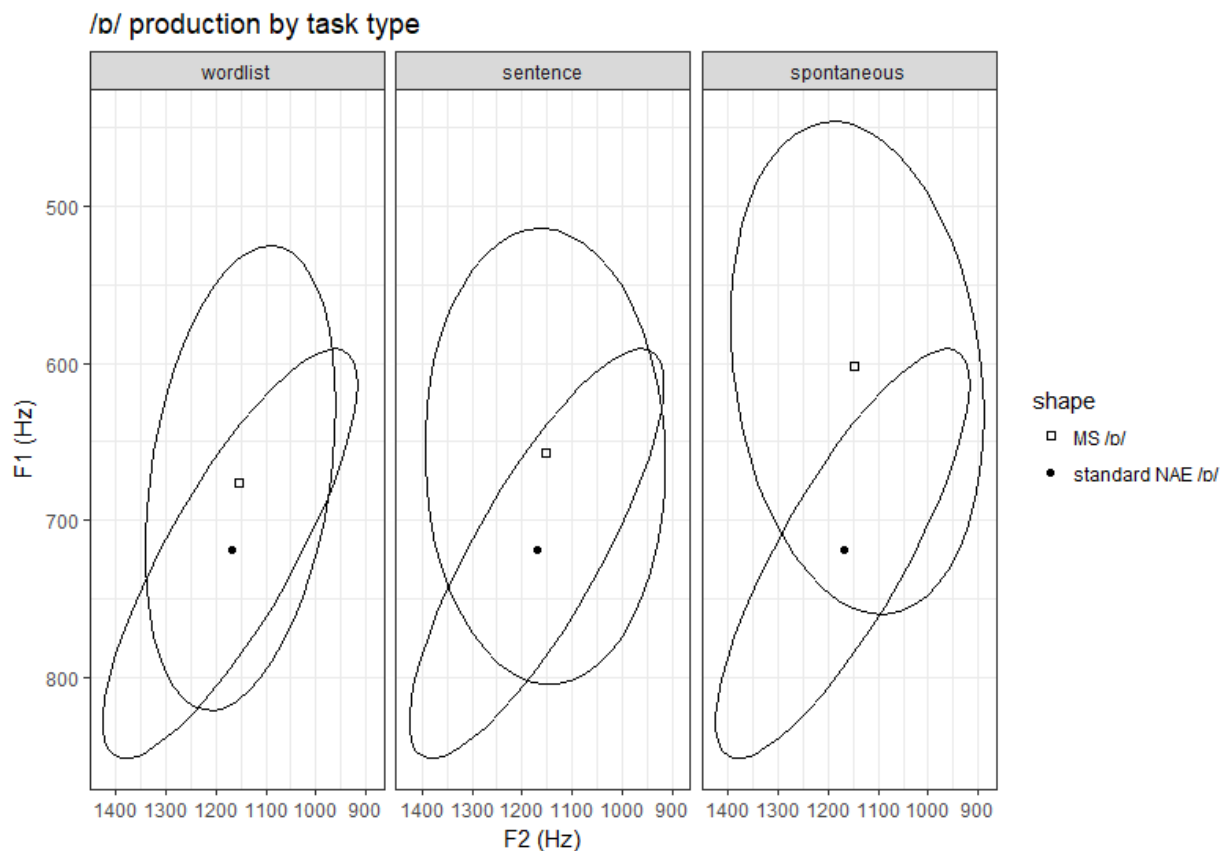


Figure 4.3 M.S.’s Acoustic productions of /ɒ/

Across task type we see that the space of M.S.’s /ɒ/ “moves up”, i.e., F1 on average seems to be steadily decreasing with task (recall from Section 2.1 and 2.5 that F1 is inversely correlated with vowel height). Mean F1 decreases from 676 Hz in the wordlist task, to 657 Hz in the sentences, to 602 Hz in the spontaneous task. M.S.’s /ɒ/ productions become acoustically less similar to the NAE standard as task formality decreases, but this acoustic difference was imperceptible.

Similar SHG Vowels

M.S.’s acquisition of /ɒ/ is interesting in how it presents a near-parallel situation to that of M.S.’s acquisition of /æ/, but with a difference that was likely key. Like how /æ/ is between /a/ and /ɛ/, /ɒ/ lies between two sounds from M.S.’s SHG vowel space: /a/ and

/ɔ/. Unlike the situation with /æ/, neither of these two sounds is contrastive with /ɒ/ in NAE, or at least, in the NAE M.S. is most likely to encounter (Canadian English, which has merged these sounds). Thus, M.S. is able to produce [ɒ] in an acoustic manner that is able to be perceived as /ɒ/ in highly reliable way. The low back vowel space of Canadian English does not have multiple competing rounded vowels, and therefore M.S. is able to produce rounded vowels that are perceived as /ɒ/ even with a degree of acoustic variance from the NAE standard. Because of the lack of competing vowels, M.S. would be able to take his SHG /ɔ/ and only need to alter its productions slightly in order for them to be acceptable as [ɒ]. The fact that variance still exists is evident in the way in which productions of [ɒ] begin to grow distant as task formality decreases (or, alternatively, grow closer as task formality increases), separating from the range of the standards and becoming more mid-vowel ([ɔ])-like.

4.4 Substitution of NAE /u:/ with SHG /u/

The high back vowel NAE /u:/ was of interest in the production study, as literature supported that this sound is phonetically different from the SHG /u:/ equivalent, which is more peripheral, gravitating towards the edge of the vowel space, whereas NAE /u:/ gravitates more towards the centre of the vowel space. Because these two /u:/s were classified as *similar* by the SLM, acquisition was predicted to be difficult, which is why a total of M.S.'s 103 productions of /u:/ were acoustically measured and analyzed.

M.S.'s production of /u:/ was split between two predictable allophones, a high central rounded vowel [ɯ], always appearing after palatal consonants and sometimes after alveolar consonants, and a high back rounded vowel [u] appearing elsewhere. While

together M.S.'s /u/ occupied the general vowel space of NAE /u/, when the allophonic variation is taken into account, a clear pattern of accuracy emerges. The table below details the variation:

Table 4.4 M.S.'s productions of /u/

Target	Production	Wordlist (T1)	Sentence (T2)	Semi-spontaneous (T3)	Total
	<i>Total</i>	30	41	32	103
/u/	[ɯ]	6 (20%)	20 (49%)	15 (47%)	41 (40%)
	[u]	24 (80%)	21 (51%)	17 (53%)	62 (60%)

Tasks

NAE /u/ can vary between back [u] and central [ɯ], and in Canadian English it is increasingly being centralized to [ɯ], with rate of centralization increasing as speaker age decreases (Boberg, 2008). M.S. produces both of these: in the wordlist task M.S. produced 6 out of 30 tokens as on-target [ɯ], an accuracy of 20%. In the sentence task his accuracy increased to 49%, 20 out of 41 tokens, and in the semi-spontaneous task it remained stable at 47%, 15 out of 32 tokens. Overall accuracy was 40%.

Acoustically, M.S.'s overall productions of /u/ completely overlap with the NAE standards collected in Kent & Read (2002). However, when broken down, M.S. produces two allophones that are acoustically stable (i.e., do not greatly vary between tasks): a central [ɯ] and a further back [u]. The below plot outlines these productions. Like with /ɒ/ ellipses surround productions of /u/ by M.S. and the NAE standards, and in this plot the two allophones have their own ellipses, with the dashed line for [ɯ] and the dotted line for [u].

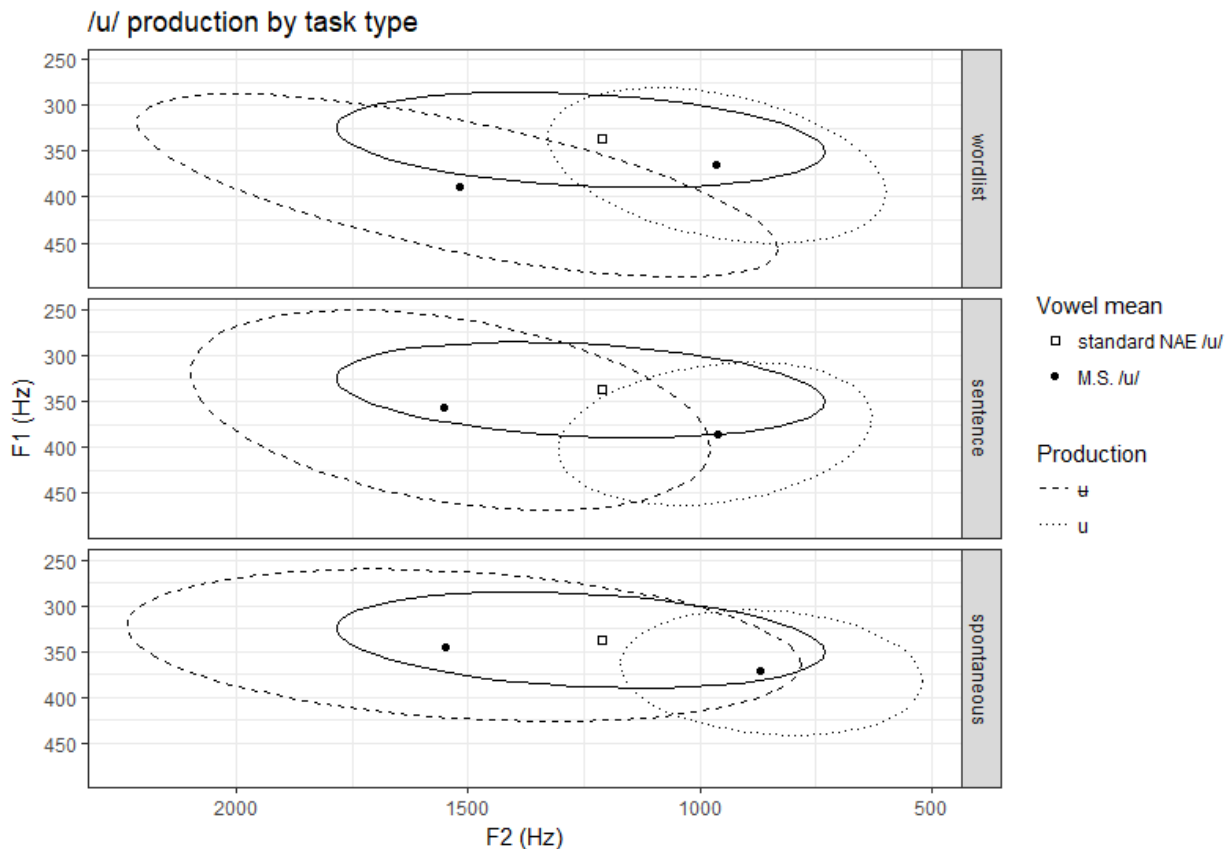


Figure 4.4 M.S.'s Acoustic productions of /u/

Acoustically the two production allophones are relatively stable across the task types. The mean of the standards lies about directly between the two, with a mean F2 of 1211 Hz. M.S.'s [ʊ] has a mean F2 of 1517 Hz, 1554 Hz, and 1548 Hz across wordlist, sentence, and semi-spontaneous tasks, respectively. M.S.'s [u] has a mean F2 of 965 Hz, 962 Hz, and 870 Hz, in the same three tasks, respectively. A Welch's 2-sample *t*-test confirmed that F2 values between M.S.'s [u] and [ʊ] productions were significantly different from each other in all tasks, with $p < .001$ for wordlist, sentence, and semi-spontaneous tasks. While acoustically M.S.'s /u/s are near evenly equidistant from the standards, M.S.'s [ʊ] is the one that is on-target for Canadian English. M.S.'s [u] is far back enough to correspond with SHG [u], and is the variant that is perceptually "off".

Therefore, of these two allophones, M.S. is producing a [u] that is no different in quality from his L1 [u], and a [ʊ] that is much more on-target and L2-like.

Syllable Context and Co-articulation

This split, however, is not due to an acquisition of the different phonetic properties of Canadian English /u/ from SHG /u/, but rather is purely the result of co-articulation.

M.S. produces on-target /u/ in a very specific environment: when preceded by a palatal or an alveolar consonant. This co-articulatory effect was only seen with palatals in the wordlist task, e.g., with [njʊ], /nju/ *new*, or [fjʊm], /fjum/ *fume*, but was seen with both palatals and alveolars in the sentence and semi-spontaneous tasks, e.g. with [tʊ], /tu/ *to*.

Of all the tokens containing /u/ that were preceded by a palatal (24 instances out of 103 instances) he was 92% on target. For the alveolar consonants preceding /u/ (38 instances out of 103 instances) M.S. was on target in 50% of all cases. For all other consonants preceding /u/ (41 instances out of 103 instances) he was never on target, which overall suggested that the accuracy of his /u/ productions was highly dependent on the consonant preceding /u/.

4.5 Devoicing of /b, d, g, v, ð, z, ʒ, dʒ/ in coda position

The voicing contrast of obstruents in coda position was predicted to be difficult for native speakers of German, as supported by existing literature. German devoices underlying voiced obstruents in coda position. Because of the systematicity of this phonological process, it was hypothesized that M.S. would produce syllable final voiceless obstruents in lieu of voiced obstruents. NAE maintains a voicing contrast between voiced and voiceless obstruents in coda position, as explained in Chapter 2.

M.S. attempted 546 underlying voiced obstruents in coda position across the three tasks, e.g., in /kɒʊv/ *cove*, /bæd/ *bad*, or /leɪdʒ/ *ledge*, but his accuracy in producing these obstruents as actually voiced varied considerably across task type as well as consonant type, with devoicing being very common, as in [kɒ:f], [bæt], or [leɪʃ], respectively. The table below shows M.S.'s accuracy for /b/, /d/, /g/, /v/, /z/, and /dʒ/ (/ð/ and /ʒ/ were not produced) across wordlist, sentence, and semi-spontaneous production tasks.

Table 4.5 M.S.'s productions of underlyingly voiced coda obstruents

Target	Production	Wordlist (T1)	Sentences (T2)	Semi-spontaneous (T3)	Total
/b/	<i>Total</i>	22	21	22	65
	[b]	0 (0%)	9 (43%)	9 (41%)	18 (28%)
	[p]	15 (68%)	11 (52%)	13 (59%)	39 (60%)
	[p ^h]	3 (14%)	0 (0%)	0 (0%)	3 (4%)
	[p̄]	4 (18%)	1 (5%)	0 (0%)	5 (8%)
/d/	<i>Total</i>	27	50	42	119
	[d]	0 (0%)	30 (60%)	31 (74%)	61 (51%)
	[t]	10 (37%)	18 (36%)	10 (24%)	38 (32%)
	[t ^h]	9 (33%)	2 (4%)	1 (2%)	12 (10%)
	[t̄]	8 (30%)	0 (0%)	0 (0%)	8 (7%)
/g/	<i>Total</i>	22	16	39	77
	[g]	1 (5%)	3 (19%)	30 (77%)	34 (44%)
	[k]	8 (36%)	9 (56%)	6 (15%)	23 (30%)
	[k ^h]	11 (50%)	3 (19%)	2 (5%)	16 (21%)
	[k̄]	2 (9%)	1 (6%)	1 (3%)	4 (5%)
/v/	<i>Total</i>	25	23	22	70
	[v]	0 (0%)	1 (4%)	1 (5%)	2 (3%)
	[f]	25 (100%)	22 (96%)	21 (95%)	68 (97%)
/z/	<i>Total</i>	24	90	75	189
	[z]	0 (0%)	12 (13%)	11 (15%)	23 (12%)
	[s]	24 (100%)	80 (87%)	64 (85%)	168 (88%)
/dʒ/	<i>Total</i>	10	6	3	19
	[dʒ]	0 (0%)	1 (17%)	1 (33%)	2 (11%)
	[tʃ]	10 (100%)	5 (83%)	2 (66%)	17 (89%)
Deletions		3	2	2	7
Total		133	208	205	546

Task 1

In the wordlist reading task, M.S. produced 133 instances of underlyingly voiced obstruents in coda position, out of which 1 was on-target, for an overall accuracy of 0.8%. This singular on-target production was of [g], in the word *executive*. Save for three instances

of deletion of underlying /d/, all other underlyingly voiced obstruents were produced as voiceless in all productions, with stops involving a mix of released, aspirated, and unreleased variants, and fricatives and affricates realised as voiceless.

Task 2

For the sentence production task, M.S. produced 208 instances of underlyingly voiced obstruents in coda position, out of which 56 were on-target, for an overall accuracy of 27%. This accuracy was higher for stops and the affricate than it was for fricatives. /b/ had an accuracy of 43%, /d/ of 60%, though /g/ of only 19% and 17% for /dʒ/. Accuracy for fricatives was 13% for /z/ and 4% for /v/. Save for two instances of deletion of /d/, all other productions were voiceless, with stops being again produced as a variety of plain, aspirated, and unreleased voiceless stops.

Task 3

In the semi-spontaneous production task, M.S. produced 205 instances of underlyingly voiced obstruents in coda position, out of which 83 were on-target, for an overall accuracy of 40%. As with the sentence production task this accuracy was greater for stops than it was for fricatives, with /b/ having an accuracy of 41%, /d/ of 74%, and /g/ of 77%, while /v/ had an accuracy of 5% and /z/ of 15%. Along with two instances of deletion of /d/, all other underlyingly voiced obstruents were produced as voiceless, with a variety of plain, aspirated, and unreleased voiceless stops.

Across Task Types

Across all tasks combined, M.S. produced 546 instances of underlyingly voiced obstruents, of which 140 were on-target, making his overall accuracy score across all tasks 26%. His most accurate segment was /d/, with an overall accuracy of 51%, followed

by /g/ at 44%, /b/ at 28%, /z/ at 12%, /dʒ/ at 11%, and his least accurate was /v/ at 3%.

While all accuracy is near zero for all segments in the wordlist task, the obstruents differ in how accuracy increases with task type. Stops increase in accuracy more from wordlist to sentence to semi-spontaneous tasks than fricatives or affricates. /b/ increases from 0% in the wordlist to remaining stable at 43% and 41% in the sentence and the semi-spontaneous tasks, /d/ increases from 0% to 60% to 74%, and /g/ from 5% to 19% to 77%; all increases were of 40%~50% or more, from wordlist to semi-spontaneous tasks. Meanwhile, /z/ increases from 0% to 13% and remains stable at 15%, /dʒ/ increases to 17% and then to 33%, while /v/ only increases to 4% and then remains stable at 5%.

This increase in accuracy across tasks is also visible when comparing words which appeared in both the wordlist task and the sentence task. Of 81 words which appeared in both, 23 were produced accurately in the sentence task but not in the wordlist task, 56 were produced inaccurately in both, and one was produced accurately in the wordlist task but not the sentence task. All the segments that M.S. produced accurately in the sentence task were stops.

Phonological environment and syllable contact

The main factor affecting overall increasing accuracy across tasks was environment. When produced before silence in the wordlist, 132 of 133 instances of underlyingly voiced obstruents were produced as voiceless. However, the other two tasks often involved obstruents produced before other speech sounds as well as silence. When the specific context is broken down, a pattern emerges in the sentence and semi-spontaneous tasks in terms of effects of the following segment, shown in Table 4.6 below. On the left

side are the various environments which followed an underlyingly voiced obstruent, and the columns break down by task the production of voiced or devoiced obstruents.

Table 4.6 Syllabic Environment Conditioning M.S.'s Obstruent Coda Voicing

Following environment	Sentence (T2)		Semi-spontaneous (T3)	
	Voiced obstruent	Devoiced obstruent	Voiced obstruent	Devoiced obstruent
Silence (utterance final)	5 (16%)	26 (84%)	0 (0%)	8 (100%)
Voiced consonant	32 (43%)	42 (57%)	42 (81%)	10 (19%)
Voiceless consonant	12 (20%)	47 (80%)	17 (26%)	48 (74%)
Vowel	12 (29%)	30 (71%)	25 (32%)	53 (68%)

Overall, unsurprisingly, devoiced obstruents outnumber voiced obstruents across all categories save one (Task 3 obstruents preceding voiced consonants). Looking at post-obstruent environments, obstruents were more likely to be voiced when they preceded a voiced consonant than when they preceded silence, a voiceless consonant, or a vowel. This environment accounts for 52% of voiced obstruent productions in the sentence task and 49% in the semi-spontaneous task. In the sentence task the environments of a following voiceless consonant and a following vowel (i.e., potential resyllabification) account for 19% and 21% respectively, with on-target utterance final productions being the remaining 8%. In the semi-spontaneous task, a following voiceless consonant and a following vowel account for the remaining 20% and 31% respectively.

In both tasks the presence of following speech sounds seems to be the main cause of the increased accuracy in production. This is especially evident from the fact a

following voiced consonant or vowel accounts for 73% and 71% of the on-target voiced productions in the sentence and semi-spontaneous tasks, respectively. These specific environments make it easier for M.S. to produce the obstruents as voiced, by allowing a general voicing environment to ‘help’ the voiced obstruent retain voicing in production. In some cases, the presence of a following vowel may even mean that M.S. resyllabifies the coda voiced obstruent to onset position of a syllable, ‘saving’ it from being devoiced.

Phonetic vs. Phonological Contrast

The main phonological contrast at work here is that of voiced ~ voiceless obstruents. This is a contrast present in both NAE and SHG. However, SHG does not have a contrast in obstruent voicing in syllable coda position. Given the discussion above on syllable context and contact as well as the evidence from the wordlist condition, M.S. can be said to not have phonologically acquired the voiced ~ voiceless distinction in NAE coda position. While he can produce voiced codas, they are conditioned by the surrounding environment and not indicative of the acquisition of the NAE contrast, especially seen through contrast with the near-zero on-target production rate in Task 1 and in Task 2 and 3 when utterance final.

Buried in the overall rates of accuracy is the variation in accuracy by different obstruent type. As noted above, stops are voiced at a considerably higher rate than fricatives. /b/, /d/, and /g/ were voiced overall at a rate of 28%, 51%, and 44% respectively, while /v/ and /z/ were only voiced at a rate of 3% and 12% respectively.

Orthographic Effects

Orthography had little influence on on-target voicing. In the wordlist task did not seem to (positively) affect M.S.’s production, as even those tokens that had perfect phoneme-

grapheme correspondence in coda-position like *grab* / __b/, *pig* / __g/, or *quiz* / __z/ were systematically devoiced in all but one word: *executive*. In the sentence task it likewise had no effect, as voiced stops are always written with a corresponding letter (i.e., /b/ with , /d/ with <d>, etc.), so there was no chance for a voiced stop to be written with a character that usually would represent a voiceless stop to interfere with voicing. Voiced fricatives, on the other hand, were written with both a letter corresponding to either a voiced or a voiceless fricative (e.g., /v/ with <f> as in *of* or <v> as in *cove*), but this did not affect on-target voicing.

4.6 Aspiration of /p, t, k/ in coda position

Unlike voiced coda obstruents, the voiceless stops /p, t, k/ do not present a phonemic issue to the native speaker of German, but rather a phonetic one. SHG has aspirated coda stops, as opposed to NAE, which only has aspirated stops in onset position. Because NAE [p, t, k] and SHG [p^h, t^h, k^h] are not contrastive, they can be regarded as *similar* sounds because they are allophones in both SHG and NAE, with the only difference in which allophone is used in syllable-final position. Acquisition of the phonetic difference would be difficult according to the SLM, which is why it was hypothesized that M.S.'s tokens containing these NAE coda stops may exhibit aspiration, carried over from his native language SHG.

In regards to accuracy scores, measuring these non-contrasting coda stops, the results summarized below could provide a measure of relative “accentedness”, if there was aspiration present:

Table 4.7 M.S.'s productions of Coda Voiceless Stops

Target	Production	Wordlist (T1)	Sentences (T2)	Semi-spontaneous (T3)	Total
/p/	<i>Total</i>	11	6	2	19
	[p̄]	3	5	0	8
	[p]	7	1	2	10
	Accuracy:	91%	100%	100%	95%
	[p ^h]	1 (9%)	0 (0%)	0 (0%)	1 (5%)
/t/	<i>Total</i>	36	56	17	109
	[t̄]	11	31	3	45
	[t]	9	19	3	31
	Accuracy:	55%	89%	35%	70%
	[t ^h]	16 (45%)	6 (11%)	11 (65%)	33 (30%)
/k/	<i>Total</i>	19	32	10	61
	[k̄]	0	6	1	7
	[k]	10	14	6	30
	Accuracy	53%	63%	70%	61%
	[k ^h]	9 (47%)	12 (37%)	3 (30%)	24 (39%)
Total		66	94	29	189

Task 1

In the wordlist task M.S. produced 10 out of 11 /p/s as either unreleased or released, an accuracy of 91%. 20 out of 36 /t/s were on-target, an accuracy of 56%, and 10 of 19 /k/s were on-target, an accuracy of 53%. In all instances, off-target productions were [p^h], [t^h], or [k^h].

Task 2

In the sentence task 100% of /p/s were on-target, 89% (50 out of 56) /t/s were on-target, and 63% of /k/s (20 of 32) were on-target.

Task 3

In the semi-spontaneous task, 100% of /p/s were on-target, only 35% of /t/s (6 of 17), and 70% of /k/s were on-target (7 of 10).

Across Tasks

Overall, M.S.'s accuracy was 69% across all tasks and segments. /p/ is always highly on-target, with a consistent accuracy over 90%. /k/ is fairly consistently middling, increasing with task type from 53% to 63% to 70% accuracy. /t/ varies wildly, jumping from 56% to 89 to 35% accuracy with task type. No single task was better than the others. The wordlist task featured the lowest accuracy for /p/ and /k/, but not for /t/. The semi-spontaneous task featured the highest accuracy for /p/ and /k/, but not for /t/.

4.7 Interdentals

The NAE interdental fricatives /θ/ and /ð/ were reported to be problematic for native speakers of German, as they are absent from the SHG phonetic inventory altogether. These sounds are well researched and reported by existing literature and considered *new* sounds to the native speaker of German, with the perceptually closest sounds being the alveolar fricatives /s/ and /z/. This substitution has been frequently reported in studies involving SHG speakers with English L2, which is why the NAE interdentals were also of interest in this study.

M.S. produced a majority of interdental fricatives as on-target in all tasks. His favoured substitution for /θ/ is [s], and for /ð/ is both [z] and [d]. The below table shows the breakdown of production across tasks.

Table 4.8 M.S.s productions of Interdentals

Target	Production	Wordlist (T1)	Sentences (T2)	Semi - spontaneous (T3)	Total
/θ/	<i>Total</i>	23	28	24	75
	[θ]	18 (78%)	14 (50%)	18 (78%)	50 (67%)
	[s]	4 (17%)	13 (46%)	6 (22%)	23 (31%)
	[t]	0 (0%)	0 (0%)	0 (0%)	0 (0%)
	[z]	0 (0%)	1 (4%)	0 (0%)	1 (1%)
	[tʰ]	1 (4%)	0 (0%)	0 (0%)	1 (1%)
/ð/	<i>Total</i>	4	17	26	47
	[ð]	4 (100%)	12 (70%)	16 (62%)	32 (68%)
	[z]	0 (0%)	1 (6%)	6 (23%)	7 (15%)
	[d]	0 (0%)	3 (18%)	3 (12%)	6 (13%)
	[s]	0 (0%)	1 (6%)	1 (3%)	2 (4%)
Total		27	45	50	122

Task 1

In the wordlist production task M.S. produced 18 of 23 /θ/ as on-target, an accuracy of 78%. He substituted with [s] four times and [tʰ] once. He produced four instances of /ð/, all of which were on-target, an accuracy of 100%.

Task 2

In the sentence production task, M.S. produced 28 instances of /θ/, of which 14 were on-target, an accuracy of 50%. He substituted with [s] 13 times and with [z] once. He produced 17 instances of /ð/, of which 12 were on-target, an accuracy of 71%. He substituted with [d] three times, and with [z] and [s] once each.

Task 3

In the spontaneous production task, M.S. produced 23 instances of /θ/, of which 18 were on-target, an accuracy of 78%. He substituted the remainder with [s] six times. M.S.

produced 16 on-target productions of /ð/, an accuracy of 62%. The substitutions were six [z], three [d], and one [s].

Across Tasks

Across task type M.S. had no clear trend in accuracy for /θ/, other than it was 50% or over in all tasks. It was 78% in Task 1, 50% in Task 2, and 78% in Task 3. Across the tasks his favoured substitution remained [s], accounting for an overwhelming majority of substitutions in all three. For /ð/ the trend was different. Accuracy somewhat decreased from 71% to 64% from Task 2 to the Task 3, and was indeed at 100% for Task 1, but there were only four productions total in Task 1, so the extreme accuracy in this task may not be indicative of how he might have performed with more tokens to produce.

Additionally, substitutions were varied more than for /θ/. Both [z] and [d] were used, with [z] appearing once in the sentence task and five times in the semi-spontaneous task, and [d] appearing three times in both.

Syllabic context

Broken down by syllable context, a clear trend is visible in M.S.'s production of /θ/ and /ð/

Table 4.9 Syllabic Context Conditioning M.S.'s Interdental Productions ("O-T" = On-target)

Segment	Environment	Wordlist (T1)		Sentences (T2)		Semi-spontaneous (T3)	
		O-T	Sub.	O-T	Sub.	O-T	Sub.
/θ/	Onset	7	0	6	0	18	0
	/θɪ/ onset	0	2	0	3	0	1
	Coda	11	3	7	11	0	5
/ð/	Onset	4	0	12	5	16	10
	Coda	0	0	0	0	0	0

/θ/ is 100% on-target across all tasks when in onset position, save for when it is part of the onset cluster /θɪ/, in which case it is never on-target, consistently substituted with [s]. In coda /θ/ is on-target at 79% in the wordlist task, 61% in the sentence task, but 0% in the semi-spontaneous task. In contrast to /θ/, /ð/ never appears in coda position, and has more variable onset-position accuracy. /ð/ is on-target 100% in the wordlist task, 71% in the sentence task, and 62% in the semi-spontaneous task. Overall, there was no consistent trend for either phoneme across task types. There was a trend of decreasing accuracy for /θ/ in coda position and /ð/ in onset position across task types, but /θ/ did not decrease in accuracy in onset position, being either always on-target when the only consonant in an onset, or always off target when in a /θɪ/ cluster.

4.8 Epenthetic [t] / Non-reduction of /t/

This section on the non-reduction of /t/ emerged from the production data as a separate phenomenon that has not been reported in existing literature. From the production data, it emerged that M.S. was producing phonetic tri-consonantal clusters containing <t> orthographically, such as the tokens *transcription*, *picture*, *facts*. In the native speaker C.E. productions, no such [t] could be detected as in [tɹænskɹɪpʃn̩], [pɪkʃə], [fæks]. Before diving into the analysis, one has to point out that there are two different phenomena at hand: the NAE words ending in {-tion}, such as *description*, *caption*, or *transcription* have a <t> grapheme, but this <t> is never pronounced. For other tokens containing <t> orthographically, the native speaker reduced clusters such as in the NAE words *picture* or *facts*: [pɪkʃə], [fæks], respectively. M.S., however, treated these two different phenomena equally, which is why they were analyzed together. Due to

sufficient data across tasks, I decided to include this observation and report epenthetic [t] in the table below:

Table 4.10 M.S.'s productions of /CtC/ clusters

Underlying	Production	Wordlist (T1)	Sentence (T2)	Semi-spontaneous (T3)	Total
	<i>Total</i>	23	24	21	68
/CtC/	[CC]	0 (0%)	0 (0%)	16 (76%)	16 (24%)
	[CtC]	23 (100%)	24 (100%)	5 (24%)	52 (76%)

In order to justify the motivation of the analysis the table below includes a selection of the tokens for which C.E. did not epenthesize [t] in {-tion} words, as M.S. did, and the productions for which C.E. reduced the underlying tri-consonantal cluster, and M.S. did not:

Table 4.11 Selection of tokens from the wordlist containing epenthetic [t] for M.S.

	token	C.E. production	M.S. production
epenthesis of [t]	transcription	[tɹænskɹɪpʃn]	[tɹænskɹɪpʃn̩]
	description	[dɪskɹɪpʃn]	[dɪskɹɪpʃn̩]
	election	[ɪlɛkʃn]	[ɪlɛktʃn̩]
	exception	[ɪksɛpʃn]	[ɪksɛptʃn̩]
non-reduction of [t]	picture	[pɪkʃə]	[pɪktʃə]
	lecture	[lɛkʃə]	[lɛktʃə]
	punctual	[pʌŋkʃʊəl]	[pʌŋktʃʊəl]
	structure	[stɹʌktʃə]	[stɹʌktʃə]
	facts	[fæks]	[fæktʃs]
	gift shop	[gɪftʃɒp]	[gɪftʃɒp]

Task 1

In the wordlist task, M.S. produced 23 instances of tokens in which <t> was present orthographically, either in words ending in <-tion>, like *election* or *exception*, or in [t̃] in *punctual* or *picture* at a syllable boundary, or [ts̃] in plural forms ending in <-ts>, such as *facts* or *tracts*. M.S. produced affricates [ts̃] and [t̃ʃ] for all of these tokens, so either those that had underlying /t/ as well as those that had {-tion} morphemes and no underlying [t]. M.S. produced sequences of stop/fricative + affricate, as in [fækt̃s̃] *facts*, but C.E. produced reduced versions that were only stop/fricative + fricative clusters as in [fæks̃], *facts*. As C.E. evidenced, NAE will reduce obstruent + obstruent clusters. And while C.E. reduced these triconsonantal (CĈC) clusters to acceptable stop + fricative clusters, M.S. did not. In fact, M.S. fails to reduce these CĈC clusters in 23 out of 23 cases. Therefore, his accuracy score for CĈC cluster reduction is 0% on the wordlist task. In all of 23 instances in which cluster reduction, or no [t] should occur, he forms stop + stop + fricative clusters, which are realized as a stop + affricate cluster. A typical instance of this pattern would be:

Table 4.12 Examples of /Ct̃C/ in NAE by C.E. and M.S.

	C.E.		M.S.	
<function>	[fʌŋk.f̃ŋ]	CVCC.CC	[fʌŋk.t̃ʃ̃ŋ]	CVCC.CCC
<facts>	[fæks̃]	CVCC	[fækt̃s̃]	CVCCC
<gift shop>	[gɪf.f̃ɒp]	CVC.CVC	[gɪft̃.f̃ɒp]	CVCC.CVC

Task 2

In the sentence reading task, M.S. produced 24 instances with [t]-epenthesis, either due to failure to reduce the $C\widehat{C}C$ cluster, or due to epenthesis of [t] into {-tion} words, all of which were reduced or not articulated by C.E. His accuracy here was also 0%.

Task 3

In the semi-spontaneous production task, M.S. produced 21 instances of epenthetic [t] for which either $C\widehat{C}C$ cluster reduction was expected, or where there was no underlying /t/. M.S. successfully reduced clusters in 16 cases here, resulting in an accuracy score of 76% overall. However, as this was a semi-spontaneous production task, production was not controlled and so all but three $C\widehat{C}C$ cluster tokens were with the word *actually* [ɛkʃəli], which he frequently used as a discourse marker. In total, he produced 18 tokens of *actually*, one instance of *pictured*, one instance of *aspects*, and one instance of *tourists*. Out of these latter three tokens, he only reduced the $C\widehat{C}C$ cluster in <tourists> [tuvi:s:], but not in <aspects>, [æspɛkʃts], and <pictured> [pɪktʃæɪt]. Moreover, while he successfully reduced the $C\widehat{C}C$ cluster in <actually> 15 times, there were three instances in which he did not.

Morphophonology

Generally, there were four different types of clusters in which M.S. epenthesized [t] or did not reduce underlying /t/. The clusters all shared the same phonological environment (detailed above), but the morphological environments differed. These morphological environments were: Latin based words ending in <-tion> which only orthographically contain a stop <ɾ>, such as *distraction* {distract} {-tion} or *transcription* {transcript} {-tion}, which in NAE is not articulated at all; words ending in stop+stop clusters followed

by the plural morpheme or third person singular morpheme {s}, e.g. *facts* {fact} {s} or *affects* {affect} {s}; and stop+stop clusters in coda-position followed by an obstruent in the onset of the following syllable, e.g. *gift shop* {gift} {shop}. For the spontaneous production task, M.S also produced instances of *actually*, which underlyingly also has a stop+affricate sequence. Moreover, there is one last group of tokens that had the same phonemic environment in which we find non-reduction, but morphologically, it is not to be grouped with the abovementioned ones, because those instances occurred at morpheme boundaries. In words like *capture* {capture}, *punctual* {punctual}, and *actually* the phonological environment is morpheme-internal, e.g., {actual} {ly}. German shares a similar morphology to English in that it has {-tion} Latin words in its lexical inventory. It also has plural morpheme {s}. Therefore, this section is closely linked to that of orthography and cognates explained below.

Cognates and Orthographic Effects

The cognate status of these tokens also plays a role. When looking at the 0% accuracy score in the wordlist and sentence production task, it stands out that a relatively high number of the words have direct cognates in German, such as NAE *construction* and SHG *Konstruktion* or NAE *instruction* and SHG *Instruktion*, due to the shared genealogy coming from Latin, so both, English and German, share the same Latin derived suffixes for some nouns.

In relation to the cognate status of some of the tokens appearing on the wordlist, it cannot be excluded that orthographic effects also play a role. The orthography of German is less opaque than that of English, meaning that German orthography more closely

reflects pronunciation than English orthography. Thus in orthography as well as in pronunciation, German has an underlying /t/ that is articulated in words ending in {-tion}. Moreover, German does not reduce triconsonantal clusters with stop+affricates, like in *packt's* 'pack it' [pakts̩] or *Aktien* [aktsiən]. German orthography could therefore support the non-reduction in English clusters as well.

A strong indicator of the influence of cognate status and orthography affecting M.S.'s productions of these English tokens is the low accuracy score in the wordlist and sentence tasks, in which he never reduces these clusters. These tasks involved M.S. reading written words presented to him, and therefore there was a direct link between what he saw, influenced by orthography, and what he produced. Furthermore, there was an increase in accuracy in the semi-spontaneous task when there was no orthographic input. However, one has to be aware that the only reason his accuracy score was higher in this task is due to his frequent use of the term *actually*. Out of the 18 productions, he successfully reduced the cluster 16 times. But it would be a tenuous claim to say that this indicates any substantial increase in accuracy, because the term was repeated over and over. This suggests that M.S. has acquired and stored the term *actually* with cluster reduction. Given that *actually* is a high frequency word in colloquial speech of NAE speakers, one can assume here that M.S. has learned how to use the specific term with cluster reduction as a result of frequently hearing it in an English speaking environment (Bybee, 2002), in which he has lived for many years, rather than retrieving its orthography or relying on a potential cognate (SHG *aktuell* = NAE *topical*). But this reduction process for this very term seems to be incomplete, as he does not consistently reduce the cluster (only in 16 out of 18 times).

Consequently, one can say that cluster reduction is not salient for M.S. in English and has not been acquired, as underlying /t/ is retained in virtually all instances, or epenthesized in {-tion} words, likely due to orthographic effects or cognate status of these words in SHG. With respect to the factors of orthography and cognate status, their effects seem to strongly influence M.S.'s pronunciation and result in transfer of these structures from his L1.

4.9 Substitution of NAE [ɫ] with SHG [l]

In SHG, there is only one phonemic /l/, that is commonly referred to as 'light' l, which is an alveolar lateral approximant, whereas in Standard British English, there are generally two different /l/s, which occur in complementary distribution. 'Light' [l] would occur in syllable onsets, whereas 'dark' [ɫ] would occur in syllable codas. However, in NAE, /l/ is realized as dark [ɫ] in all positions (Celce-Murcia et al., 1996 ; Davenport & Hannah, 2013), meaning that there is no purely alveolar lateral, but only a velarized alveolar lateral approximant [ɫ]. Conclusively, while SHG and NAE share /l/, their realizations in their languages differ. The analysis of the Canadian English speaker (C.E.) that served as a control for this experiment also confirmed the hypothesis that there is no [l] in NAE English. Due to their phonetic similarity and their allophonic status in British English, light /l/ and dark l [ɫ] can be considered *similar* sounds, according to the SLM, which would predict that phonetic acquisition is *difficult*, because NAE [ɫ] is a near equivalent to SHG /l/ and they do not phonemically contrast.

The table below shows a total of 300 NAE target /l/ productions by M.S. across tasks, reporting the substitutions and deletions that occurred and providing accuracy scores for each task type, respectively.

Table 4.13 M.S.'s productions of /l/

Target	Production	Wordlist (T1)	Sentence (T2)	Semi-spontaneous (T3)	Total
	<i>Total</i>	80	138	82	300
/l/	[ɫ]	24 (30%)	35 (25%)	16 (19%)	75 (25%)
	[l]	56 (70%)	102 (74%)	63 (77%)	221 (74%)
	Deleted	0 (0%)	1 (1%)	3 (4%)	4 (1%)

Task 1

In the wordlist task, M.S. produced 80 instances of underlying /l/, out of which 24 were on-target productions (30%). The most common substitution type, and the only substitution type, that M.S. produced were 56 instances of [l] (70%). In total, M.S. achieved an accuracy score of 30% for the wordlist task.

Task 2

For the sentence production task, M.S. produced 138 instances of underlying /l/, out of which 32 were on-target productions (25%). M.S.'s most common substitution type was [l], which he substituted in 102 instances (74%). There was one instance of deletion (<1%). In total, M.S. achieved an accuracy score of 25% for the sentence reading task.

Task 3

In the spontaneous production task, M.S. produced 83 instances of underlying /l/ out of which 16 were on-target productions (19%). The most common substitution type was [l],

which he substituted 63 times (77%). The spontaneous production task also exhibited 3 deletions (4%), leaving M.S. with an overall accuracy score of 20%.

Overall, across tasks, M.S. achieved a cumulative accuracy score of 24% with 73 on-target productions. The most common substitution across tasks was [l], which occurred 223 times out of 300 tokens (75%), 4 productions were deletions across tasks, making up 1%.

Syllable context

There is a relatively clear trend to be found when examining the phonological environment in which [t] occurs. In the wordlist task, 45 out of 80 tokens had /l/ occurring in onset position, whereas 35 /l/ tokens occurred in codas. Out of all of these 80 tokens, M.S.'s got 24 on-target productions. Only 4 of those (17%) were onset productions of [t], whereas the other 20 tokens (83%) appeared in coda position. Therefore, only 11% of all onset tokens were on-target productions. There were 35 codas of /l/ in the wordlist, out of which 20 were on-target [t] productions (57%).

For the sentences, 77 out of 138 tokens appeared in onset position, and 61 in coda. Out of all 138 tokens, M.S. had 35 on-target productions. 7 of those occurred in onset position and 28 occurring in coda.

And finally, for the spontaneous task, a similar trend could be observed. Out of 82 tokens for the spontaneous production task 16 were on-target productions overall. 3 of those occurred in onset whereas the other 13 were in coda.

Table 4.14 M.S.'s productions of /l/ by Syllabic context

	Wordlist (80 tokens)		Sentences (138 tokens)		Spontaneous (82 tokens)	
	Onsets 45 (56%)	Codas 35 (44%)	Onsets 77 (56%)	Codas 61 (44%)	Onsets 48 (59%)	Codas 34 (41%)
On-target	4 (11%)	20 (57%)	7 (9%)	28 (46%)	3 (6%)	13 (38%)
Others	41 [l] (89%)	15 [l] (43%)	70 [l] (91%)	33 [l ; -] (54%)	45 [l] (94%)	20 [l ; -] (62%)
Mean accuracy	Onsets across tasks 15/170 8%			Codas across tasks 61/130 50%		

Vowel Environment

Since the velarized l (dark [ɫ]) shares some features with high back vowels (Jones, 1966 ; Wells, 1982 ; Whitley, 2002), examining the vowel environment preceding and following underlying /l/ was crucial. However, the tokens exhibited a very high variety of different front, central, and back vowels preceding and following [ɫ], which made it impossible to establish a pattern. The most striking factor seemed syllable context alone.

4.10 Substitution of /w/ with /v/

In this section of the results, the NAE /w/-/v/ contrast which were produced by M.S. had been reported to be problematic for native speakers of SHG by the existing literature (Swan and Smith, 2001; Chamson, 2016). SHG has only the labiodental fricative /v/ in its phonetic inventory, but lacks the labiovelar glide /w/. Despite some phonological similarities of /w/ and /v/, as explained below, /w/ was classified to be a *new* sound to the native speaker of SHG, as the phonetic distance between /w/ and /v/ is large enough for /v/ not to be considered a near equivalent of /w/, according to the SLM. However, research reports that this /w/ sound of NAE is often “overgeneralized” by native speakers

of German, where they substitute NAE /w/ with SHG /v/, or even NAE /v/ with /w/ (Chamson, 2016), suggesting that these two sounds do not form two phonemically contrastive categories to the native speaker of SHG. Therefore, the production data obtained from a total 147 instances of NAE /w/ was analyzed.

The table below shows all instances of underlying /w/ by production task and M.S.'s realizations thereof (on-target productions, substitutions, and deletions) whose mean accuracy score is calculated according to task type.

Table 4.15 M.S.'s productions of /w/

Target	Production	Wordlist (T1)	Sentence (T2)	Semi-spontaneous (T3)	Total
	<i>Total</i>	23	54	68	147
/w/	[w]	3 (13%)	11 (20%)	11 (16%)	25 (17%)
	[v]	14 (61%)	35 (65%)	45 (66%)	94 (64%)
	[ʋ]	4 (18%)	7 (13%)	12 (18%)	23 (16%)
	[ʉ]	1 (4%)	1 (2%)	0 (0%)	1 (1%)
	[v̄w]	1 (4%)	0 (0%)	0 (0%)	1 (1%)
	Deleted	0 (0%)	0 (0%)	0 (0%)	1 (1%)

Task 1

In the wordlist task, M.S. produced 23 instances of underlying /w/ out of which 3 were on target productions (13%). The most common substitution type was [v] in 14 cases (61%), followed by 4 instances of [ʋ] (17%), as well as one instance each of [ʉ] (4.5%) and [v̄w] (4.5%). Overall, M.S.'s accuracy score for this task was 13%.

Task 2

For the sentence production task, M.S. produced 54 instances of /w/ out of which 11 were on-target productions (20%). The most common substitution type was [v] which he used 35 times (65%), followed by seven instances of [ʊ] (13%) and one deletion (2%). For the sentence production task, M.S. therefore produced an accuracy score of 20%.

Task 3

In the semi-spontaneous production task, M.S. produced 68 instances of underlying /w/, out of which 11 were on-target productions (16%). His preferred substitution type was again [v], which he used 45 times (66%), followed by 12 instances of [ʊ] (18%), leaving him with an accuracy score of 16%.

Across all tasks combined, M.S. produced 147 instances of /w/, out of which 25 were on-target productions of [w] (17%). The most common substitution type was [v] in 94 of all cases (64%), followed by [ʊ], which he substituted in 23 instances (18%). Deletion (<1%), [ɥ] (<1%) substitution, and [v̄w] (<1%) substitution each occurred once across all tasks.

Across all three tasks, M.S.'s accuracy score remains relatively stable with the lowest accuracy score (13%) occurring in the wordlist task, followed by the semi-spontaneous task (16%), and reaching the highest accuracy score in the sentence task (20%), resulting in an overall accuracy score of 17% across tasks. M.S.'s favoured substitution type across tasks is /v/, which he substituted 94 times out of 147 (64%). The percentage for this substitution type remains relatively stable across tasks as well, so a strong preference by M.S. for [v] productions in lieu of /w/ could be found. Moreover, the analysis of the data showed an almost equal amount of [ʊ] substitutions across tasks (16%). Therefore, one can conclude that M.S. has a strong preference for [v], but when

he DOES produce a glide, there is an even chance for it to be either a target-like [w], or a [v]. Interestingly enough, [v] is a glide that is absent from both the inventories of NAE and SHG.

Phonetic Dissimilarity

According to the accuracy score across all tasks, which never exceeds 20%, M.S. has not phonologically acquired /w/. However, as the data show, he has a 17% accuracy score for [w] productions, so he can articulate /w/, as in phonetically produce it. The low accuracy score and the high amount of substitutions, especially [v], suggests that M.S. has not established a distinct category for /w/, which he maps onto his L1 category of /v/.

Moreover, the glide substitution [v] provides evidence that he is aware of the manner of articulation of /w/ and attempts to produce on-target /w/, but fails to fully get the place of articulation right, which is labiodental for /v/, as opposed to labio-velar for /w/.

One can establish that, phonetically, /w/ and /v/ differ considerably in place and manner of articulation, with [v] being a fricative and [w] a glide; they also differ in lingual articulation, [v] having no lingual articulation and [w] having raising of the tongue body close to the velum. Their similarity comes in compression of the lips, but here they likewise differ, as [v] has spread lips.

Phonological Similarity

In order to establish the phonological similarity of /w/ and /v/ to M.S., one needs to compare the phonological features of /v/ and /w/ as done in the table below. Labial features are in focus because this is the most salient feature M.S. appeals to, as all of his /w/ substitutions are exclusively labial sounds, and the other labial consonants are included to highlight that /w/ is the only [dorsal] sound.

Table 4.16 Comparative Features of NAE and SHG Labial sounds

Cross-comparison of all labials	/m/	/p/	/pf/	/b/	/f/	/v/	/w/	/ʋ/
consonantal	+	+	+	+	+	+	-	-
sonorant	+	-	-	-	-	-	+	+
approximant	-	-	-	-	-	-	+	+
continuant	-	-	-	-	-	+	+	+
nasal	+	-	-	-	-	-	-	-
voicing	+	-	-	+	-	+	+	+
LABIAL	✓	✓	✓	✓	✓	✓	✓	✓
DORSAL							✓	

(Davenport & Hannahs, 2010)

One can take a grounded approach to features and note that by phonological features /w/, /ʋ/, and /v/ are very similar due to the involvement of the lips and that in listening there will be voicing and a lack of nasality. Therefore, one might argue they are featurally similar, more so than any other labial sounds present in NAE or SHG. The features in the table are mentioned above and serve to show how these sounds might be similar in the minds of speakers due to shared features.

As has been established, the most salient feature that M.S. relies on when producing non-target-like /w/s are the features /v/, /ʋ/, and /w/ have in common. Comparing with /w/, one can see that /w/ has more in common with /v/ than any other labial sound **in SHG**. Apart from the difference of /w/ being an approximant, and therefore also sonorous, as well as dorsal, /v/ shares all of its other features: both /v/ and /w/ are voiced, continuants, and non-nasals, as opposed to all other labials, which are not continuant, not voiced, or nasal. In fact, to M.S., /w/ seems to be perceptually similar to

his L1 category of /v/ that he equates this bilabial glide with his native category of a labiodental fricative. The labiodental glide /v/ also appears, though it is not a sound of SHG, because it fills in the features associated with glides that /v/ lacks.

Cognate Status and Orthographic Effect

Orthography is crucial in the possible effects of cognate status, and since the <w>-words of English are of Germanic origin, they have formally similar cognates in German:

Table 4.17 NAE and SHG <w> Cognates

NAE	Phonetic transcription	SHG	phonetic transcription
world	[wɜ:ld]	Welt	[vɛlt]
wonder	[wɒndə]	Wunder	[vʊndɐ]
way	[weɪ]	Weg	[ve:k]
water	[wɔ:tə]	Wasser	[vasɐ]
whale	[weɪl]	Wal	[va:l]

The difference between <w> in both languages comes from German undergoing a sound change in the period from 1200-1400 when SHG transitioned from Middle High German to New High German. /v/ merged to /f/ becoming voiceless, and consequently /w/ became /β/ and then later /v/ to fill this category, causing /w/ to disappear from the German language altogether (Salmons, 2012). The orthography, however, did not change along with the sound change, so the new /v/ sound is still represented by <w> to this day (Chambers & Wilkie, 1970). The formal resemblance between the cognate words is relatively high (due to shared genealogy) and therefore, when M.S. reads or retrieves an English word with <w>, and consequently a /w/ sound, his German lexicon may be activated as well, retrieving a formally similar German word causing M.S. to produce English words with underlying /w/ with /v/ instead, as indicated by the table above.

Vowel environment in which /w/ Appears

Finally, when [w] is produced, it appears in a specific environment with respect to the following vowel: [w] predominates when the following vowel is a back vowel, and [v] is most substituted when the following vowel is a front vowel. The below table shows the total amount of M.S.'s [w]-productions and their substitutions appearing according to task type and vowel environment.

Table 4.18 Vowel Environment by Task conditioning /w/ productions

Vowel environment	[w]			[v]			[ʊ]			[ʉ]			Total Vowel
	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3	
[i]		1	1			5			1				8
[ɪ]	1			6	11	9	1						28
[e]		1				3		1	1				6
[ɜ]				2	5	1	1	1		1			11
[ɛ]			6	6	7	11	1	1					32
[ɑ]	2	4		2		3		1	1				13
[ʌ]		2	24		3	5		2	6				42
[ɔ]		1	3		1	1							6
[aɪ]			1		5	4	1		2				13
total productions by task	3	9	35	16	32	42	4	6	11	1	0	0	

While across task types no obvious correlation of task type with vowel environment and accuracy/substitution could be established, the raw numbers suggest that M.S. produces more on-target [w] when the following vowel is [+back]. The wordlist task (T1) had too few numbers to reflect a trend, but the sentence production task (T2) and the semi-spontaneous production task (T3) show that there is a correlation between the place of articulation of vowels and production accuracy of /w/. This correlation can be more easily seen when we collapse across task type, as done in the table below, which shows percentage of productions relative to vowel environment.

Table 4.19 Summary of Vowel Environment conditioning /w/ productions

Following Vowel	[w]	[v]	[ʊ]	[ɯ]	Total
[i]	25.00%	62.50%	12.50%	0.00%	100.00%
[ɪ]	3.57%	92.86%	3.57%	0.00%	100.00%
[e]	16.67%	50.00%	33.33%	0.00%	100.00%
[ɜ]	0.00%	72.73%	18.18%	9.09%	100.00%
[ɛ]	18.75%	75.00%	6.25%	0.00%	100.00%
Average front	13%	71%	15%	2%	100%
[ɑ]	46.15%	38.46%	15.38%	0.00%	100.00%
[ʌ]	61.90%	19.05%	19.05%	0.00%	100.00%
[ɔ]	66.67%	33.33%	0.00%	0.00%	100.00%
[aɪ]	7.69%	69.23%	23.08%	0.00%	100.00%
Average back	58%	30%	11%	0%	100%

From this table, one can see that across all three task types, M.S. produces 58% of on-target [w] when it is followed by a back vowel, as opposed to only 13% of on-target productions when [w] is followed by a front vowel. The opposite pattern can be established for the substitution type [v], which predominantly occurs after front vowels at a rate of 71%, as opposed to only 30% of [v] substitutions occurring after back vowels. In order to confirm the robustness of this trend, a Welch's 2-sample t-test was done on the distributions of the two productions as they relate to vowel front/backness. The rate of [w] appearing before [-back] vowels was compared to the rate of [w] appearing before [+back] vowels; the same was done for [v] in the same environments. The trends described are statistically significant for effect of environment. The p-value for on-target [w] predominantly occurring before [+back] vowels was $p < .01$ and for substitution [v] predominately occurring after [-back] vowels the p-value was also $< .01$. Moreover, it could be established that there was no correlation between the [v] substitution and vowel

environment, as the difference between [v] preceding front and back vowels showed no significant difference. For front vowels, M.S. substituted [v] at a rate of 15%, and for back vowels he substituted [v] at a rate of 11%. A t-test shows that the difference between the rates was not statistically significant, as $p=.69$.

It is likely that M.S. produces more on-target instances of [w] when preceding a back vowel because of the featural similarities in place of articulation between back vowels and /w/. Where /w/ is [+high] and [+back], both /ʊ/ and /ʌ/ are [+back], and /ʊ/ even shares [+high] with /w/. Meanwhile, the front vowels are all [-back] and therefore more dissimilar to /w/ than the back vowels, thus M.S.'s higher accuracy with back vowels.

Hypercorrection of /v/ to /w/

The literature review (see Section 2.5.2) discussed the possibility of M.S. hypercorrecting NAE /v/ to [w], and so as part of the investigation of NAE /w/ this topic was investigated. However, little effect of hypercorrection was found, as seen in the table below:

Table 4.20 M.S.'s productions of /v/

Target	Production	Wordlist (T1)	Sentence (T2)	Semi-spontaneous (T3)	Total
	<i>Total</i>	15	31	17	63
/v/	[v]	14 (93%)	27 (87%)	16 (94%)	57 (91%)
	[w]	0 (0%)	3 (10%)	1 (6%)	4 (6%)
	[ʊ]	1 (7%)	1 (3%)	0 (0%)	2 (3%)

In the wordlist production task, M.S. produced 15 tokens with /v/ in syllable onset out of which 14 were on-target productions of /v/ (93%). One token was a [ʊ]-substitution

that appeared in the word *viral* (7%). In the sentence production task, M.S. produced 31 tokens of /v/ in syllable-onsets, out of which 27 were on-target productions (87%), whereas three [w] substitutions occurred (10%) and one substitution of [v] (3%). In the semi-spontaneous production task, M.S. produced 17 instances of syllable-onset /v/, out of which 16 were on-target productions (94%). In one instance, he substituted [w] for /v/ (6%). Across all tasks, M.S. produced 57 out of 63 on-target instances of /v/ in syllable-onset, resulting in an overall accuracy score of 91%. Only 6 tokens were off-target productions, representing [w] (6%) and [v] (3%), respectively.

Because of his strong preference for [v] over [w], even for /w/ tokens, hypercorrection of /v/ to [w] is not a feature in M.S.'s speech.

4.11 Prevocalic /ɹ/

Prevocalic /ɹ/ was also a segment of interest for this study, as SHG lacks this alveolar liquid entirely, the closest equivalent being the SHG uvular fricative /ʁ/. However, the difference between these two rhotics is relatively large when comparing their phonological features. Therefore, the SLM predicted that phonetic acquisition of this *new* sound to native speakers of German would be *easy* due to its phonetic dissimilarity from the SHG counterpart. Although it is predicted to be perceived as different from any SHG sound by the native speaker of SHG, existing literature supports that production of this sound will usually be **gradually** acquired for native speakers of German due to its complex articulation (lip rounding, tongue bunching, approximant), which called for analysis of M.S.'s production data for this *new* consonant.

Overall, M.S. produced prevocalic /ɪ/ as on-target at about chance rate. He varied between on-target [ɪ], as in [gɹeɪvjɑ:d] *graveyard*, labio-dental glide [v] as in [kvɪm] *cream*, and total deletion, as in [fɪm] *from*. The table below summarizes productions across task type. There was not a single substitution with the SHG counterpart [ɜ]:

Table 4.21 M.S.'s productions of Prevocalic /ɪ/

Target	Production	Wordlist (T1)	Sentence (T2)	Semi-spontaneous (T3)	Total
	<i>Total</i>	97	100	87	284
/ɪ/	[ɪ]	58 (60%)	43 (43%)	36 (41%)	137 (48%)
	[v]	39 (40%)	51 (51%)	33 (38%)	123 (43%)
	Deleted	0 (0%)	6 (6%)	18 (21%)	24 (9%)

Task 1

In task 1 M.S. produced 58 of 97 /ɪ/ tokens as on-target, an accuracy of 60%. The remaining 39 tokens were produced as labio-dental glide [v].

Task 2

In task 2, 43 of the 100 tokens were on-target (43% accuracy), and 51 of the remain tokens were produced as [v], with six more deleted.

Task 3

In Task 3 36 of 87 tokens (41% accuracy) were produced as on-target [ɪ], and 33 were produced as [v] with 18 deleted.

Across Tasks

Accuracy declined from Task 1, at 60% to Task 2, at 43%, and stabilized in Task 3 at 41%. Overall, [v] was the favoured substitution, peaking in proportion in the sentence task, accounting for 40%, 51%, and 38% of total /ɪ/ tokens across tasks. The sentence task also saw the appearance of a new trend for M.S.'s /ɪ/: deletion. While only 6% in the

sentence task, in the semi-spontaneous task 21% of /ɹ/ tokens were deleted in M.S.'s speech. Thus, an interesting trend appears: while on-target productions stabilize with decreasing task formality, substitution with [v] begins decreasing, perhaps being replaced by wholesale deletion, but of course, given the nature of the two tasks, this is impossible to know for sure.

When comparing words that appear in both the wordlist task and the sentence task, no obvious and robust trend appears. 15 words that appeared in both tasks were on-target in both, 15 words in both tasks were off-target in both, 13 declined from wordlist to sentences, and 6 improved from wordlist to sentences. Additionally, following vowel environment had no obvious correlation with accuracy, as neither did orthography (NAE is /ɹ/ always written with <r>).

M.S. substitutes instead with the same labio-dental approximant that he uses for NAE /w/: [v]. It is likely that M.S. is picking up on the lip rounding of [ɹ], failing to substitute with a [w], and instead using a [v]. /l/ is included as it is the other English liquid. The feature chart below shows the similarities of /ɹ/, /w/, /v/, and /l/:

Table 4.22 Comparative Feature chart for /ɹ/

Cross-comparison of all labials	/ɹ/	/w/	/ʊ/	/l/
consonantal	+	-	-	+
sonorant	+	+	+	+
approximant	-	+	+	+
continuant	+	+	+	
nasal	-	-	-	-
lateral	-	-	-	+
voicing	+	+	+	+
LABIAL	✓	✓	✓	
DORSAL	✓	✓		(✓)

(Davenport & Hannahs, 2010)

With the exception of /l/, we see that /w/ is the closest NAE sound in terms of features to /ɹ/, and M.S. often uses the glide closest to /w/, [ʊ], as a substitution for /w/ (see Section 4.11).

This substitution is not an uncommon pattern for L2 (James, 1988) or even L1 acquisition of /ɹ/, where it is often one of the last sounds acquired (Dodd, Holm, & Hua, 2003), usually moving through stages of varying [w]-likeness before finally being acquired (James 1988).

Overall, M.S. performs at chance in accuracy of on-target /ɹ/ production, and this accuracy is generally freely varying, and so he cannot be said to have acquired it.

4.12 Substitution of /ɹ/ in Postvocalic Position

Although NAE has /ɹ/ in all syllable positions, and section 4.7 reports M.S.'s productions of NAE prevocalic /ɹ/, this section had to be treated separately, due to a phonological process of SHG that vocalizes the SHG uvular /ʁ/ to the vowel [ɐ] (dark schwa), which is

an allophonic variant of /ʁ/, that exclusively occurs after vowels. A similar process can be observed in British English, as explained in section 2.5.3, which, however, does not exist in NAE, where /ɪ/ is maintained in all positions regardless of syllable context. Since the vocalization of SHG /ʁ/ has been anecdotally described in the existing literature (Grantham O'Brien & Fagan, 2016 ; König & Gast, 2009), this segment felt worth investigating in M.S.'s production data, in which he produced a total of 221 tokens of NAE /ɪ/ in postvocalic position.

The table below reports M.S.'s produced instances of /ɪ/ across tasks, giving his accuracy scores and the substitution types occurring for /ɪ/, respectively.

Table 4.23 M.S.'s productions of Postvocalic /ɪ/

Target	Production	Wordlist (T1)	Sentence (T2)	Semi-spontaneous (T3)	Total
	<i>Total</i>	60	99	62	221
/ɪ/	[ɪ]	19 (32%)	44 (45%)	11 (18%)	74 (33%)
	[ɐ]	20 (33%)	44 (45%)	10 (16%)	74 (33%)
	[ə]	8 (13%)	1 (1%)	18 (29%)	27 (12%)
	[ɐ̃]	3 (5%)	4 (4%)	0 (0%)	7 (3%)
	[ɜ̃]	1 (2%)	0 (0%)	0 (0%)	1 (2%)
	[ɔ̃]	1 (2%)	0 (0%)	11 (18%)	12 (5%)
	Deleted	8 (13%)	6 (6%)	12 (19%)	26 (12%)

Task 1

In the word reading task, M.S. produced 60 instances of /ɪ/ in postvocalic position. 19 of those were on-target productions (32%), whereas the most common substitution type was a dark schwa [ɐ] – an allophonic variant of the German uvular fricative /ʁ/ -- which

occurred 20 times (33%). This substitution type was followed by [ə], which M.S. used 8 times (13%). Additionally, M.S. deleted post-vocalic /ɪ/ in 5 cases, all of which occurred when /ɪ/ was preceded by the vowel /a/, such as in [ka:d] for /kaɪd/, ‘card’, or [gɹeɪvjɑ:d] for /gɹeɪvjɑɪd/, *graveyard* (8%). In those instances, underlying /ɪ/ could be argued for due to an increase in vowel duration from [a] to [a:]. The same pattern applies to /ɜ:/, for which M.S. deleted the r-colouring by reducing to [ɜ] in three tokens (5%). In two instances, M.S. produced /ɜ/ and /ə/ (stressed and unstressed rhotic vowels), respectively (3%), with a weakened r-colouring, so [ɪ] was barely audible, albeit “present”, which is why these tokens were acoustically measured to see if /ə/ was present. For example, in the token [bɛrɛɪ] NAE *better*, a higher F3 of /ə/ was acoustically measured than was for a fully realized [ɪ], which was indicative of M.S.’s tongue not “bunching” as much as one would expect for /ə/.

Within the group of weakened r-coloured schwas, M.S. also produced three [ɐ] substitutions (5%). The cumulative mean of all substitutions and deletions leaves M.S. with an accuracy score of 32% for post-vocalic /ɪ/.

Task 2

The sentence reading task contained 99 tokens of postvocalic /ɪ/, 46 of which occurred in the wordlist as well. Out of these 99 instances, 44 productions were on target (45%). Another 44 instances of postvocalic /ɪ/ were substituted with [ɐ] (45%). In six cases, there was a deletion of /ɪ/ when preceded by the vowel /a/. In those cases, M.S. lengthened /a/ to [a:] (6%). Additionally, M.S. produced [ɐ] with weakened r-colouring in four instances (4%). There was only one [ə] substitution (1%).

Task 3

In the semi-spontaneous production task, M.S. produced 62 instances of postvocalic /ɪ/ out of which 11 were on target (18%)⁶. The most common substitution type was [ə] in 18 cases (29%). 11 times, M.S. produced the vowel (schwa) of a token with weakened r-colouring that was barely audible, albeit present (18%). There were 10 [ɐ] substitutions (16%). In eight instances, M.S. deleted postvocalic /ɪ/ following other vowels than /a/ (13%). Additionally, there were four deletions of /ɪ/ when preceded by /a/ (6%), all of which M.S. lengthened from /a/ to /a:/. As opposed to the wordlist and sentence production task, these deletions did not only occur when /ɪ/ was preceded by vowel /a/, but also when /ɪ/ followed other vowels.

The overall accuracy across all task types is 34%, which means that M.S. does not successfully preserve /ɪ/ in postvocalic position as is typical of NAE.

Across Task Types

As established, 46 of the tokens appearing in the wordlist also appeared in the sentences. Out of those 46 items M.S. produced, 23 (50%) were equally bad exemplars of postvocalic /ɪ/ in the wordlist as well as the sentence task. Seven tokens were equally good exemplars of postvocalic /ɪ/ in the wordlist and sentence task (15%). In six instances, M.S. performed better on the wordlist task than the sentence task (13%). However, in another ten instances, M.S.'s produced more on-target exemplars of postvocalic /ɪ/ in the sentence task than in the wordlist (21%). As a conclusion, M.S.'s performance slightly improved in the sentence production task, even though 65% of the

⁶ It is important to know that M.S. produced a high number of the same token, such as 'university', which may have affected the accuracy score.

identical tokens that occurred in the wordlist and sentences exhibited no difference in quality (better or worse) from the wordlist to the sentence task.

Phonetic vs. Phonological Contrast

As the segment /ɪ/ in postvocalic position was under examination in this group, it was important to establish whether a phonetic or phonological issue can account for M.S.'s low accuracy score across tasks. In order to do so, M.S.'s productions of /ɪ/ in prevocalic position were considered as well (see previous Section 4.11). Throughout all three production tasks, M.S. produced 206 instances of /ɪ/ in prevocalic position. His accuracy score for /ɪ/ in prevocalic position was 48%, which means that he has not phonetically acquired a phone that is foreign to the inventory of his native language, though phonemically NAE and SHG inventories both have a rhotic phoneme. Though he is not always on-target with [ɹ] prevocalically, he does not vocalize it prevocalically, so he has successfully acquired the NAE pattern (which in this case is the same as the SHG pattern). This observation leads to conclusion that M.S. has a low accuracy score for /ɪ/ in postvocalic position due to a phonological issue: he is applying r-vocalization that occurs in SHG to NAE.

The possibility of this was noted in the section on /ɪ/ in Chapter 2.5.3. SHG vocalizes uvular fricative /ʁ/ to [ɐ] when /ʁ/ occurs in coda position. For a similar pattern from English, see non-rhotic varieties of English, e.g., RP, or Australian or New Zealand English, where /ɪ/ vocalizes to [ə] when in coda position. In NAE however, /ɪ/ is realized as [ɹ] in all syllabic environments. M.S. is therefore likely using the r-vocalizing pattern from his L1 in his L2, where instead of /ʁ/ being realized as [ɐ] it is /ɪ/ being realized as [ɐ] or [ə]. Further evidence comes from his preferred substitution type: he favours [ɐ]

over every other substitution type across tasks, with the exception of the semi-spontaneous production task. Overall, across all three tasks, 74 out of 220 instances were [ɐ] substitutions (33%), the sentence reading task exhibiting the highest number of [ɐ] substitutions (45%). While the third production task exhibits more /ə/ substitutions than [ɐ] substitutions, this may be due to the speech rate. Overall, M.S.'s strong preference for the dark schwa [ɐ] suggests that this is likely L1 transfer. Additionally, the large number of tokens that are cognates in English and German supports this claim. If NAE <*bear*> is <*Bär*> in German (semantically identical and phonetically similar), and M.S. produces [beɐ] in his native language, this could heavily influence his L2 production.

Cognate Status

The English language bears an extensive number of Germanic words with /ɪ/. For example: NAE *bear* and SHG *Bär*, NAE *wonder* and SHG *Wunder*, or NAE *winner* and SHG *Gewinner*. Not only do they have orthographic similarities, but they are semantically identical as well (hence translations of the German item are not necessary in this case). Moreover, English and German are morphologically similar when it comes to inflections as well. For example, in German, {-er} is a common morpheme, either as a plural inflection ('Kinder' *children*, 'Räder' *wheels*), a comparative morpheme ('besser' *better*, 'schneller' *faster*), or nominalized verbs to indicate what a person is doing ('Gewinner' *winner*, 'Banker' *banker*). In English, these morphological structures are almost identical to those of German with the exception of the SHG plural morpheme {-er}, which English does not have. This results in a high frequency of /ɪ/ or /ʌ/ occurring in postvocalic position in both languages. The reason this morphological aspect is brought into the discussion under the paragraph of Cognate Status, is because English and

German do not only share such morphological similarities, making a high number of words cognates or almost identical, but because most of the English tokens occurring in the wordlist and sentences exhibit those morphological traits. Thus, the items used in those tasks are semantically and phonetically similar due to their shared etymology, which is also evident in their orthography (direct cognates). In the wordlist alone, over 30% of the tokens are cognates, such as: *polar, bear, hair, water, morning, and number*. In the sentence task, 43% of the tokens are cognates in English and German. In the spontaneous production task, 41% were cognates. These numbers only represent those tokens that are direct cognates of English and German, not necessarily taking into account the ones that share the same morphology. It could be observed that M.S.'s performance on cognates was 27% accuracy for cognates in the wordlist task. Thus, out of 22 tokens that are cognates, only six were on target. For the sentence task, his accuracy on cognates improved to 44%, as 19 out of 43 cognates were on-target productions. For the spontaneous task, M.S. produced 26 cognates, out of which 5 were on target (19%). Therefore, a pattern could be established that suggests that *if* M.S.'s productions are not target-like, this inaccuracy commonly occurs tokens that are direct cognates. Therefore, it can be assumed that especially for direct cognates, M.S. activates his L1 lexicon (Dijkstra et al., 2010), applying the familiar phonological structures from his native language. Compared with the on-target productions for non-cognates, throughout all tasks, by tendency, M.S. performed better on non-cognates than cognates. Especially in the wordlist task, M.S.'s produced twice as many on target productions on non-cognates than cognates. This evens out a little more in the sentence production task, in which M.S. produces on-target non-cognates at an accuracy rate that exceeds that of

cognates by approximately $\frac{1}{4}$. In the semi-spontaneous production task, 11 productions were on target, out of which 5 were cognates and 6 were non-cognates, but this may well be due to M.S. often repeating tokens.

Orthographic Effects

In addition to their cognate status, orthographic effects may contribute to M.S.'s production of these tokens. Orthographically, when finding <r> in postvocalic position, the German reader knows that underlying /ʁ/ is produced as /r/ instead. This can be seen particularly at the example of NAE *hard* and NAE *market*, both of which are cognates in German. Underlyingly, in NAE, we have /ɪ/ following the low back vowel /a/. In German, the /a/ in cognates *hart* and *Markt*, is lengthened to [a:] to indicate there is an underlying /ʁ/: /hɑʁt/ to [ha:t] and /maʁkt/ to [ma:kt]. M.S. applied this very pattern in the English tokens *heart*, *market*, *partner*, *garlic*, *card* and *are* by lengthening the vowel [a] to [a:] to indicate that there is underlying /ɪ/, as he could see by the spelling of the tokens in the first two production tasks.

Stressed VS. Unstressed Syllables

Across all three task types, it is evident that M.S. has a higher accuracy score for on-target productions if postvocalic /ɪ/ occurs in stressed syllables. In the wordlist production task, 32% productions of postvocalic /ɪ/ were on-target. In the wordlist production task, out of 19 on-target productions of postvocalic /ɪ/, 11 occurred in stressed syllables (58%) as opposed to 8 in unstressed syllables (42%). The same applies to the semi-spontaneous production, in which 18% of all postvocalic /ɪ/ productions were on target: 73% of those on-target productions occurred in stressed syllables, as opposed to 27% on-target productions occurring in unstressed syllables. The most interesting

observation could be made in the sentence task, in which M.S. achieved the highest accuracy score: 45% on-target productions (44/98). Out of these on-target productions, 75% occurred in stressed syllables, as opposed to 25% in unstressed syllables. However, taking all three tasks into account, M.S. has the tendency to perform better on postvocalic /ɪ/ occurring in stressed syllables. In all tasks combined, M.S. produced postvocalic /ɪ/ accurately in stressed syllables at a rate of 48%, as opposed to an 38% accuracy rate in unstressed syllables. However, while there appeared to be a trend in M.S.'s production data, a Pearson's Chi-squared test found no significant relation between a stressed syllable and on-target production of postvocalic /ɪ/ ($p=.36$).

It is possible that for /ə/, M.S. has more difficulties producing the r-colouring due to this phoneme occurring in unstressed position and therefore – with English being a stress-timed language – being less salient than /ɜ:/, which is the stressed equivalent with a longer duration than /ə/ (Celce-Murcia, Brinton, & Goodwin, 1996).

4.13 Results of the Native Speaker Judgments

After the wordlist and sentence production data had been collected from M.S. and C.E., two native speaker listeners of NAE from the Department of Linguistics at the University of Victoria listened to a selected amount of productions from the sentences both speakers M.S. and C.E. had produced. There were three different components that were measured: accentedness, intelligibility, and comprehensibility. Measurement for comprehensibility and accentedness were done with adopting the Likert- 'goodness rating' scale from 1-9 (1 = very easy/no foreign accent ; 9 = extremely difficult/very thick foreign accent) (Munro and Derwing, 1995). For example, M.S. produced the following utterance:

Table 4.24 Summary of Native Speaker Judgements

	errors affecting intelligibility (mean of NSJs transcription errors)		comprehensibility ratings (mean among both NSJs 1-9 Likert scale)		accentedness ratings (mean among both NSJs 1-9 Likert scale)	
	MS	CE	MS	CE	MS	CE
Utterance 1	2.5	0	6.5	1.5	6	1.5
Utterance 2	.5	0	1.5	1.5	4.5	1
Utterance 3	0	0	2	1	5.5	1.5
Utterance 4	.5	0	2	1	4.5	1.5
Utterance 5	.5	0	3.5	1	5.5	2
Utterance 6	0	0	2	1	5	1
Utterance 7	0	0	2	2.5	5	2
Utterance 8	.5	0	3.5	1	6.5	1
Utterance 9	1.5	0	5	1.5	6.5	1
	6 errors in total	0 errors in total	3.1	1.3	5.4	1.4

Transcriptions errors by both NSJs were counted and averaged, resulting in an average error score of 6 across all utterances for M.S., whereas the NSJs had no transcription errors in C.E.'s utterances. For comprehensibility, the mean score across all utterances is 3.1 for M.S. (relatively easy to understand) and 1.3 for C.E. (1= very easy to understand). The accentedness rating showed the largest discrepancy between M.S. and C.E.'s utterances, for which M.S. received a mean score of 5.4 (5= noticeable foreign accent) and C.E. received a mean score of 1.4 (=no foreign accent).

A Welch's 2-sample *t*-test revealed that between speakers, there was a significant difference between M.S. and C.E. for intelligibility ($p=.04$), comprehensibility ($p=.02$), and accentedness ($p<.01$). As far as the NSJs transcriptions and ratings showed there is a difference between M.S.'s and C.E.'s production. This supports the inference that the NSJs consistently perceived M.S.'s speech to be less intelligible, less comprehensible, and more accented than C.E.'s speech.

Although all three measurements suggested that M.S.'s production differed from that of C.E. in all three domains, looking at the actual numbers (scores) he received for

comprehensibility suggests that he is still relatively easy to understand (3.1 = relatively easy) despite having a foreign accent (5.4 = noticeable foreign accent). Overall, both NSJs had a positive correlation ($[r=.94]$ and $[r=.71]$) between M.S.'s intelligibility and his comprehensibility, so that his comprehensibility rating increased (elevated effort) as the number of intelligibility errors increased: as he became less intelligible, he also became less comprehensible. However, the NSJs had a weak positive/negative correlation between intelligibility and accentedness ($[r=.43]$ and $[r=-.11]$). So as accentedness increased, intelligibility errors did not. M.S.'s accentedness ratings remained stable across utterances, whereas comprehensibility and intelligibility did not. Furthermore, the NSJs also responded to the questionnaire (Murphy, 2014) about what they had noticed in particular, and their observations pertained to segmental productions to a very high extent. Both, NSJ 1 and NSJ 2 noticed M.S. substituting the NAE interdental /θ/ and /ð/ with /s/ and /z/, respectively which could be found in some of their transcriptions, in which they correctly orthographically transcribed <booth> and <three>, but mentioned in parentheses that they had heard [bu:s] *boos and [sɪi] *sri, which corresponds to findings on NAE interdental in coda-position and onset clusters /θɪ/ made in the result section 4.8 on interdentals.

NSJ 2 also transcribed the word *grammar* in Utterance 8 as *grandma* (see Result section on postvocalic-r), which made *grammar* and *grandma* homophonous in M.S.'s speech production. NSJ 1 commented: “non-rhotic *grammar*”. Also, in regards to final devoicing, NSJ 1 mentioned that the word <verbs> sounded like [vəps] in M.S.'s speech production. There were additional comments they gave on suprasegmentals, like intonation and speech rate, but those were excluded from the result section as they are not

in focus for this study. Both NSJs indicated on the questionnaire spreadsheet that even though speaker 1 (M.S.) “sounded different [from speaker 2 (C.E.)]” and “definitely not from here”, the accentedness did not interfere with the intelligibility or comprehensibility, except for the words they pointed out that they were unable to understand, which were few in number.

Chapter 5 Discussion

This chapter discusses the results for M.S.'s NAE segmental productions, reported in Chapter 4, and analyses them in light of the research questions raised in Chapter 1. First of all, the accuracy scores for all NAE segments under investigation will be reported (5.1), which will show which segments of NAE have been acquired by M.S., and which ones have not. Because vowels and consonants were analyzed differently, with acoustic analysis supplementing auditory analysis for the vowels only, as explained in section 3.4, the discussion will focus on vowels and consonants separately and discuss how a clear pattern of phonetic and phonemic differences emerged for each. Upon determining which vowels and consonants of NAE are in fact difficult to produce for M.S., I will explain what factors (intra-linguistic and extra-linguistic) may influence M.S.'s interlanguage. Following this, I will discuss the constitution and state of M.S.'s interlanguage grammar, and the notion of 'variable competence'. Then I turn to a discussion of the predictions made by the Speech Learning Model and how these predictions were largely not borne out by the results. Finally, I wrap up with a word on global accentedness/comprehensibility/intelligibility, limitations and concluding remarks.

5.1 Overall Accuracy

Table 5.1 shows the mean accuracy scores across all three tasks (wordlist, sentences, semi-spontaneous speech) for each segment of NAE considered in this study. Some segments emerged from the production data that have also not been previously addressed in research, but were worth looking into, as sufficient data was provided (/ɒ/ and /ʌ/). The threshold for accuracy was >80%, as explained in Chapter 2.1, which marks the divide

between ‘acquired’ and ‘not acquired’. Scoring was reported throughout Chapter 4 as the ‘accuracy score’ for a given segment, each of which represents a distinct phonemic category in NAE. According to accuracy measures, M.S. has acquired one segment (/ɒ/) and has not acquired the other 11 reported in Table 5.1 below:

Table 5.1 Summary of M.S.’s Results split by vowel vs. consonant

target NAE segments		acquired (accuracy score)	not acquired (accuracy score)
vowels	/æ/		X (39%)
	/ʌ/		X (65%)
	/ɒ/	√ (92%)	
	/u:/		X (40%)
consonants	voiced obstruents /b, d, g, v, z, ð, ʒ, dʒ/ in coda-position		X (26%)
	unaspirated /p, t, k/ in coda-position		X (69%)
	interdentals /ð/ and /θ/		X (67%)
	epenthetic/non-reduction [t]		X (24%)
	dark ‘l’ [ɫ]		X (24%)
	/ɹ/ in prevocalic position		X (48%)
	/ɹ/ in postvocalic position		X (33%)
	labiovelar glide /w/		X (39%)

While this table reports the overall accuracy scores that M.S. achieved for each segment of interest, the discussion of these results aims to clarify that these accuracy scores in isolation cannot give much information about M.S.’s acquisition. They do not reflect that some of the segments with low accuracy scores may not be problematic, despite non-target-like production, i.e., the production of these segments may contribute to accented speech, but they don’t necessarily impair comprehensibility or intelligibility, such as NAE vs. SHG /u/ or NAE vs. SHG /l/.

5.2 Vowels

Upon examining the vowels and their accuracy scores, only one vowel, namely /ɒ/, could be considered “acquired”, while all other vowels had variation in production from native speakers of NAE, which is why they were acoustically measured and reported as off-target productions. The productions that were off-target were off-target in different ways. Some, such as /u:/ and /ʌ/ were off-target but not in a way so as to sound like a different NAE vowel. Others, like /æ/, were. This is where a very important distinction has to be made between phonetic versus phonemic errors.

“Phonetic” is used to mean the vowels of NAE that M.S. produced which differed acoustically from NAE productions. For example, for tokens containing NAE /u:/ and /ʌ/, M.S.’s productions were off-target because he substituted them with a similar vowel from his L1 SHG. Indeed, the acoustic quality of M.S.’s /u:/ productions was that of SHG [u:] in NAE words like *mood*, *flute*, or *include*. M.S. also substituted NAE /ʌ/ with a near equivalent from SHG, which is absent from NAE as a monophthong: [a]. One could argue that these are off-target productions, meaning their acoustic properties do not match those of NAE when compared to native speakers (i.e., C.E.’s measurements and the acoustic measurements of NAE vowels from Kent & Read 2002).

The accuracy score for /æ/ across tasks was 39%, but the nature of the error was a different one from that described for /u:/ and /ʌ/. With /æ/, we are dealing with a phonemic error, which is why it stands out. Phonemic errors are errors that involve substituting one NAE sound for another phoneme, thus producing a contrasting lexical item and hindering understanding. M.S. produces NAE words like *ban*, *land*, and *bag* with [ɛ], which is the closest SHG equivalent to /æ/. The problem that occurs here is that

the vowel [ɛ] he produces occupies the space of a distinct phoneme in NAE: /ɛ/. NAE /ɛ/ and /æ/ are contrastive. This means that if one is substituted for the other, it will result in a change in meaning of the word: NAE *Ben*, *lend*, and *beg* do not have the same meaning as NAE *ban*, *land*, and *bag*. For /u/ and /ʌ/, M.S.'s productions do not cause ambiguity, as his SHG equivalents occupy the accepted phonological space of NAE /u/ and /ʌ/, even if the productions are not phonetically on-target. The table below illustrates this difference between phonetic and phonemic errors and why a phonemic error is more impactful to communication than a phonetic error:

Table 5.2 M.S.'s Phonetic/Phonemic Vowel differences

Token	M.S.'s Production	Phonetic/Phonemic
<i>mood</i>	[mu:d]	phonetic difference; does not change the meaning
<i>flood</i>	[flad]	phonetic difference, does not change the meaning despite substitution
<i>lawn</i>	[lɔn]	on target
<i>land</i>	*[lænd]	phonemic difference, substituting /æ/ with [ɛ] is problematic because /æ/ and /ɛ/ represent two distinct phonemic categories in NAE

5.3 Consonants

Given the importance of distinguishing phonetic and phonemic errors for vowels, consonants were also grouped into phonetic or phonemic error categories. For the consonants of NAE that M.S. produced, a similar pattern to the vowels could be observed. While the accuracy scores suggest that none of the consonants of NAE under investigation have been acquired to a nativelike level (i.e., passed the threshold of >80%), the distinction of phonetic and phonemic errors can be made here as well.

M.S.'s erroneous productions of 'dark' /l/, /ɹ/ in postvocalic position, non-reduction of [t], and aspiration of coda-stops /p, t, k/ were phonetic errors. While they

could account for foreign accentedness, the substitutions M.S. used for these segments were not contrastive with another phoneme of NAE. For example, M.S. substituting a SHG [l] for NAE [ɫ] would not affect the meaning of the word *fool*, as English does not contrast [l] and [ɫ]. The same applied to M.S. predominantly vocalizing his /ɹ/s in post-vocalic position by substituting [ɐ], like in *bear*, his non-reduction of [t] in consonant clusters, like with [kʰts̩] instead of [ks] for /kʰts/ in *facts*, and his aspiration of stops /p, t, k/ in coda position to [pʰ,tʰ,kʰ], like in *pack*. Again, these transfers from his L1 SHG do not change the meaning of these words, which is why they are good exemplars of his NAE that are merely accented, but do not necessarily affect comprehensibility.

M.S. also had phonemic errors for consonants. These happened for NAE sounds that are absent from the SHG phonemic inventory, where M.S. substitutes the NAE consonant what he considers an equivalent or good exemplar of the underlying NAE phoneme. For example, substitution of prevocalic-/ɹ/ with a labiodental approximant [ɹ] (a sound not in NAE) that shares some features of [ɹ] or substituting NAE [w] with [ɹ] while simultaneously treating NAE /v/ as [ɹ] or [w], which shows that there is a lack of phonemic contrast for him between /w/ and /v/. The same applies to the NAE interdental fricatives /θ/ and /ð/, which are absent from SHG, for which he favours the substitution [s] and [z], respectively. This is problematic because NAE contrasts /θ, ð/ and /s, z/, such as in the NAE words *think* and *sink*. Lastly, the voicing contrast in obstruents in coda position being absent from SHG causes M.S. to devoice all of his codas, making them ambiguous: *bag* versus *back* are indistinguishable in his speech production. The distinction between phonetic and phonemic errors in consonants is illustrated in the table below:

Table 5.3 M.S.'s Phonetic/Phonemic Consonant differences

Token	M.S.'s Production	Phonetic/Phonemic
<i>fool</i>	[fu:l]	phonetic difference; substituting underlying 'dark' l with 'light' l will not change the meaning of the word; [l] and [ɫ] are not contrastive
<i>bear</i>	[bɛɐ̯]	phonetic difference; substituting [ɪ] with [ɐ̯] will not change the meaning of the word. /ɐ̯/ and /ɪ/ are not contrastive
<i>facts</i>	[fækt̪s̪]	phonetic difference; non-reduction/epenthesis of /t/ will not change the meaning of the word; they are not contrastive
<i>pack</i>	[pæk ^h]	phonetic difference; aspiration of the coda stop /k ^h / will not change the meaning of the word, it is not contrastive with [k]
<i>rack</i>	[væk]	phonemic difference; substituting [ʊ] (can sound like [w] or [v]) for [ɪ] can change the meaning of the word, it is contrastive
<i>west</i>	[vest]	phonemic difference; substituting [w] with [v] changes the meaning of the word, /w/ and /v/ are contrastive
<i>vile</i>	[waɪl]	phonemic difference; substituting [v] with [w] changes the meaning of the word, /v/ and /w/ are contrastive
<i>think</i>	[sɪŋk]	phonemic difference; substituting [θ] with [s] changes the meaning of the word, /θ/ and /s/ are contrastive
<i>bag</i>	[bæk]	phonemic difference; substituting [g] with [k] changes the meaning of the word, /g/ and /k/ are contrastive (this applies to all voiced obstruents in coda-position)

Having established the necessity of distinguishing phonetic and phonemic errors, we can see that only five segments (1 vowel and 4 consonants) fall into the category of phonemic errors, which means that only in these instances is it crucial for M.S. to accurately produce the contrast to avoid miscommunication, decreased intelligibility and comprehensibility.

In this section I summarized the segments that had been under investigation for this study and grouped them by vowel and consonants pointing out why some substitutions M.S. phonetically (acoustically) deviate from the NAE target sounds, whereas there are also vowels and consonants of NAE that are in two phonemic contrastive categories for which M.S. only had one category.

5.4 Intra-Linguistic and Extra-Linguistic Factors

In this part of the discussion, all errors, phonetic and phonemic, will be examined and interpreted to establish as much about M.S.'s interlanguage grammar as possible. In order to do so, one will have to take the accuracy scores into account and establish which intra-linguistic factors (task type, orthography, cognate status, phonetic environment, syllable context, stress) seem to be the ones that correspond with accuracy. The pattern of how M.S. has acquired NAE segments can be further understood through looking at the extra-linguistic information obtained through the qualitative elements of the study: the interview and questionnaire. Finally, with the help of the NSJs, a big picture and global understanding of M.S.'s interlanguage will emerge.

The accuracy scores reflect whether a segment of NAE can be considered “acquired” (>80%) or “not acquired” (<80%). Therefore, a ranking of the NAE segments that M.S. has partially acquired is presented in the table below, in descending order from the highest mean accuracy score across all three tasks to the lowest, across all segments.

Table 5.4 M.S.'s Overall Accuracy of segment productions in descending order

NAE segment	/b/	[p], [t̪], [k]	/θ, ð/	/ʌ/	prevocalic /ɪ/	/u:/	/æ/	postvocalic /ɪ/	/b, d, g, v, z, ð, ʒ, dʒ/ in coda-position	non-reduction of /t/	[ɪ]	/w/ - /v/
Accuracy score	92%	69%	67%	65%	48%	40%	39%	33%	26%	24%	24%	17%
Acquired	√											

Upon comparing all segments with each other for the various intra-linguistic factors they had been examined for, the table below was generated to summarize the

factors that played a role in characterizing the segments of M.S.'s interlanguage grammar. Every factor that was found to influence production accuracy is listed in the table below with check marks [√] indicating that an intra-linguistic factor played a role for a specific segment of NAE. The dash [-] indicates that the intra-linguistic factor did not play a role for production accuracy of a segment. The dashed line between the factors orthography and cognate indicates that these two variables were merged for the discussion, as they were as they were confounded with one another, with cognates of NAE and SHG sharing similar orthography.

Table 5.5 Factors affecting M.S.'s Productions

Segment	task type	orthography	cognate	phonetic environment	syllable context ⁷	stress
/æ/	(√)*	-	-	-	-	-
/ʌ/	-	-	-	-	-	-
/ɒ/	-	-	-	-	-	-
/u:/	-	-	-	√	√	-
/b, d, g, v, z, ð, ʒ, dʒ/ in coda-position	√	-	-	-	(√)	-
/p, t, k/ in coda	-	-	-	-	(√)	-
/θ/, /ð/	-	-	-	-	√	-
epenthetic/non-reduction of [t]	√	√	√	-	-	-
[ɫ]	-	-	-	-	√	-
prevocalic /ɪ/	-	-	-	-	(√)	-
postvocalic /ɪ/	-	√	√	-	(√)	√
/v/ - /w/ contrast	-	√	√	√	-	-

“√” = factor relevant to production

“-” = factor not relevant to production

*(√) means that at first glance this factor was relevant to production accuracy, but effects were minimal (see also footnote 7).

⁷ The segments of NAE prevocalic and postvocalic /ɪ/, voicing contrast of obstruents in coda position, and coda stops /p,t,k/ did not have to be examined for the influence of syllable context on production accuracy, as these phenomena are investigated *because* they pose difficulties to native speakers of SHG in this specific syllable context to begin with.

Looking at the table we can see that no one segment is affected by all factors examined. The following sections investigate each factor and discuss how it was relevant for the production of NAE segments in M.S.'s grammar.

5.4.1 Task Type

As has been explained in the literature review, different task formality types have been investigated in various studies (Silveira, 2004; 2007 ; Dickerson, 1975), some but not all found that the more formal the task type, the greater the accuracy for the speaker. This seems unsurprising at first, because one would assume that with the orthographic input supplementing the lexical knowledge of the speaker in more formal tasks, the level of difficulty would decrease. However, when taking all the target segments that have been looked at into account, only two showed a clear (i.e., no other factor could account for improvement) correlation between task type and accuracy. Moreover, it should be said that while significant task type effects could only be found for two segments (coda voicing contrast and non-reduction of/epenthetic [t]; minimal ones for /æ/ that could be attributed to other factors), this did not mean that the most formal task (wordlist) was always the one with the highest accuracy, which contradicts the hypothesis that the more formal the task the more accurate the production (Gatbonton, 1975; Dickerson, 1975). Both of these segments, namely non-reduction and epenthesis of /t/ as well as coda voicing contrast of obstruents, were in fact worse in the more formal task types when taking the accuracy scores into account. As Tarone (1987) and Ellis (1984) claim, differences in task type performance are indicative of 'variable competence', meaning that M.S. accesses a specific grammar in one task, e.g. wordlist, causing him to receive lower accuracy scores, and a different grammar in another task, e.g. sentences, resulting

in higher accuracy scores. For both, the epenthetic [t]/non-reduction of /t/ and coda voicing contrast of obstruents, we find a significant increase in M.S.'s production accuracy from the most formal to the least formal task. However, in section 5.5 I will argue that these apparent task type effects cannot be attributed to M.S.'s grammatical competence changing, and that while his performance dramatically improves for both of these segments, it does not reflect 'variable competence' or activation of different grammars.

5.4.2 Orthography

For a number of segments, namely non-reduction/epenthesis of [t] in clusters, discrimination between [w] and [v], postvocalic [ɹ], orthography could be determined to play a crucial role. As explained above, reduction/non-epenthesis of [t] improved and reached its highest accuracy score in the semi-spontaneous task. One of the factors accounting for this would be the orthographic input of the wordlist and sentences. This would lead to hyperarticulation in the more formal tasks, a plausible explanation (Colantoni et al., 2015), given M.S.'s accuracy scores for the wordlist and sentences is 0%, and accuracy dramatically increases to 76% in the semi-spontaneous task. M.S.'s ability to reduce these clusters seems to be suppressed in the formal tasks in which he epenthesized significantly more, and could be ascribed in part to orthographic influences (e.g., *transcription*, *lecture*, *facts*, *gift shop*). As explained in Chapter 2, German orthography is rather transparent in that its spelling in that its graphemes often directly correspond to phonemes, as opposed to English, which is more opaque (e.g., NAE *transcription* [tɹænskɹɪpʃn̩] and SHG *Transkription* [tʁanskʁɪptʃjo:n]). Therefore, it is likely that in the two more formal tasks M.S. does not reduce clusters because

orthographically, there is a <t> present, leading him to assume that it should be articulated. The NAE words containing {-tion} morpheme are those that do not have any /t/ underlyingly. Thus, no matter whether these words would occur in a formal or informal speech production task, the /t/ that M.S. epenthesis does not exist phonetically for native speakers of NAE.

Orthography also plays a role with NAE /w/. Because <w> represents the sound /v/ in SHG, it can be assumed that spelling, like in NAE *water*, *wall*, or *wonder*, influences M.S.'s production of these tokens due to the SHG spelling that correlates with a different sound than in NAE. Recall that accuracy score was consistently low for /w/ across tasks. This even applied to the semi-spontaneous production, supporting the view that lexical activation still activates orthographic representations and has orthographic effects, though other effects, such as cognate status, are likely also have an effect, discussed in more detail below. Post-vocalic /ɪ/ vocalization also had likely orthographic effects with the English {-er} morpheme in words like *banker*, *seller*, *better*, which correspond to orthography of SHG equivalents as in *Bänker*, *Verkäufer*, and *besser*, the {-er} being a cue to vocalization of /ɪ/ for native speakers of SHG. This will be further discussed in the section below on cognate status. The same applies to /v/ - /w/, as there may be a tie in with cognate status as well. For a number of words M.S. encountered, like *polar* and *bear*, they have spelling nearly identical in SHG and NAE. This can explain why, when M.S. encounters a post-vocalic <r> in spelling, it triggers his L1 and the /ɪ/ is processed through his L1 phonological system, because SHG only has vocalized post-vocalic /ɪ/s, unlike fully rhotacized NAE.

In investigating the role that orthography played in influencing production accuracy, the concomitant factor of cognate status consistently arose. In these cases, M.S. could be applying the rules of orthographic grapheme-phoneme correspondence of his L1, leading to a combined orthography/cognate effect. The following section addresses this issue of cognates.

5.4.3 Cognates

This intra-linguistic factor ties in with orthography. Cognate status played a role for the same three segments for which orthography was important: non-reduction and epenthesis of [t], /v/ - /w/ contrast, and postvocalic /ɪ/ vocalization. When talking about cognates, we talk about the formal similarity, i.e., the phonological or orthographic form of these words (Carroll, 1992). Though a lot of these tokens were etymologically related, Carroll does not consider this a relevant factor because speakers do not have synchronic knowledge of lexical etymology. For example, for post-vocalic /ɪ/, M.S. employs the pattern he is familiar with from his L1, namely that SHG vocalizes /ɪ/ if it is preceded by a vowel. In the more formal tasks of wordlist and sentences, it is likely a combination of orthography and cognate status playing a role. But as this /ɪ/ vocalization is also found in the semi-spontaneous production task, without orthographic input, we can assume that cognate status in interplay with orthography plays a significant role. M.S. can access a word like NAE *bear* in his L1, which happens to be SHG *Bär*, and this would lexically/phonologically activate the similar form in the L2, which is why his L1 phonology is transferred to the L2. For non-reduction/epenthesis of /t/, cognate status may be the most important factor, as most of the tokens on the wordlist and sentences were cognates in SHG and NAE. NAE *transcription* and SHG *Transkription* are true

cognates in the sense that they present almost identically in spelling and also denote the same thing. M.S.'s accuracy scores for cognates suggested that his L1 phonology of these tokens overruled L2 cluster reduction, and triggered epenthesis, as he never successfully reduced /t/ in either the wordlist task nor the sentence task.

One outlier that boosted his accuracy score significantly in the semi-spontaneous task was NAE *actually*, which is formally a cognate with SHG *aktuell*. *Aktuell*, meaning 'topical', 'recent', does not mean the same as NAE *actually*, and so, without orthography, his productions of *actually* do not feature any L1 transfer effects: he produces [ækʃəli], not [æktʃəli]. Thus, we have to take into account here that the high accuracy of producing *actually* can be ascribed to frequency effects, as M.S. might use this word frequently in his L2 or hears this word frequently when articulated by NAE speakers.

As far as /w/ and /v/ are concerned, there were many formally similar and/or etymologically related words present throughout the tasks. For example, NAE *wonder* and SHG *Wunder* or NAE *wart* and SHG *wart* (pst. tense of *sein*, 2nd pers. Pl). There is a high rate of production of [v] substitutions for NAE /w/ in formally similar NAE – SHG cognates. However, while cognate status and orthography certainly play a role in /v/ - /w/ alternation, M.S.'s low accuracy score for underlying /w/ can be attributed to the perceived similarity of the two sounds to M.S.

Concluding, for the sections on orthography and cognate status, it is not entirely clear whether production errors can be ascribed to either cognates or orthography, as those NAE segments that were investigated in this study and for which orthographic effects were suspected also appeared in words of NAE that have a cognate in SHG.

5.4.4 Phonetic Environment

The surrounding phonetic environment had an effect on accurate production of the /w/ - /v/ contrast in particular. M.S. was most accurate when /w/ was followed by a back vowel, and while his accuracy score is never near the 80% threshold, the phonetic environment had an impact on his production accuracy, which showed that he is capable of production on-target [w] at a much higher success rate with a back vowel involved. Surprisingly, /w/ was the only segment that showed vowel environment effects, whereas the other segments for which one might assume such an effect to occur, for example /ɹ/ as another approximant of NAE, such observations could not be made.

M.S. has not acquired NAE /u/ phonetically, as his productions of /u/ are usually more retracted, and therefore more SHG-like. He does, however, at times produce NAE-like /u/, but this is conditioned by consonantal environment, as they are always on-target productions when the preceding consonant (the onset of the syllable) is a palatal consonant /j/, as in [nju], /nju/ *new*, or /tʃ/, as in [tʃuwiŋ], /tʃuwiŋ/ *chewing*, and they are often, but not always on target productions when the preceding consonant is an alveolar consonant, such as /t/ /n/ or /l/ as in [tu], /tu/ *too*. Articulatorily, alveolars and palatals are further forward in the mouth than velars, which is why we can assume that M.S.'s accurate productions of NAE /u/ are conditioned by co-articulatory effects which would pull his [u] productions closer to the centre, where NAE /u/ is articulated. However, by “on target” it can only be acknowledged that M.S. produces NAE /u/ in the range of native speakers of NAE, according to Kent and Read (2002), but not in the exact range for a NAE /u/ following palatals and alveolars.

5.4.5 Syllable Context

This section focuses on /p, t, k/ coda aspiration, prevocalic and postvocalic /ɹ/ accuracy, and /b, d, g, v, z, ð, ʒ, dʒ/ devoicing in coda position. The target sounds are, by definition, restricted by syllabic environment, being either onsets or codas. /p, t, k/ are unreleased, or at least unaspirated in NAE coda position, whereas in SHG they can be aspirated.

Prevocalic /ɹ/ only appears before vowels, hence, in onsets, and postvocalic /ɹ/, only appears after vowels, hence, in codas. /b, d, g, v, z, ð, ʒ, dʒ/ are voiced in coda position in NAE, but in SHG they are devoiced.

Two other segments showed syllable context effects: 1. interdental /θ/ and /ð/ and 2. [ʃ], syllable environment showed effects on M.S.'s production accuracy. For the interdentals /θ/ and /ð/, there was a contrast in accuracy between when they appeared in onset vs. when they appeared in coda. M.S.'s production accuracy was mostly decreased in codas and was relatively high in onsets. /θ/ was always on-target in onset position, with the exception of complex onset cluster /θr/, where /θ/ was always produced as [s], and /ð/ was on-target more than it was replaced in onset position. On the other hand, outside of the isolated productions of the wordlist, all but one coda were [s] and [z] substitutions. This is in line with the idea that the coda position of a syllable is the most marked typologically, and therefore, production in it is the most restricted (Clements, 1990; Zec, 1995). Moreover, /θ/ (and its voiced counterpart) is a segment that is typologically highly marked: an interdental obstruent, which is rare across languages (Moran et al., 2014). As this stands out for M.S. and he has phonetically acquired the interdentals /θ/ and /ð/ in onset, one could argue that the combination of acquiring a difficult marked sound in a marked syllable position is a dominant factor contributing to lower accuracy scores.

Another segment for which syllable context played a role is NAE [ɫ]. As I have argued, NAE /l/ is ‘dark’ in all positions, but some varieties of English and some scholars argue for an allophonic distinction between [l] and [ɫ], with the ‘dark’ [ɫ] occurring in syllable coda, and the ‘light’ [l] in onset. M.S.’s productions to an extent reflect this syllabic-positional distinction of light and dark /l/. SHG only has light /l/ in all positions. He was significantly more accurate in production of [ɫ] codas than [ɫ] onsets. A possible explanation for this is that if /l/ appears in a coda, like in *fool* or *Jill*, it is usually preceded by a vowel or an approximant, the most sonorous of speech sounds, and therefore the most vowel-like (vowels being the most sonorous). My hypothesis is that, it is easier to produce a dark [ɫ] if the vowel has already been produced, dark /l/ being more vowel-like (Mielke, Baker, & Archangeli, 2016). Though one may expect it to only be easier to produce dark [ɫ] after back vowels, an allophonic pattern that does happen in Kirghiz, Turkish, Kharia, Alabaman, and Norwegian, it is also just as likely for /l/ to be always dark in coda position, as happens in Southern British English, Dutch, Mundari, Taos, and Brazilian Portuguese (Mielke et al., 2016). This cross-linguistic tendency indicates that there are likely universal articulatory factors that influence M.S.’s more accurate production of dark [ɫ] in codas.

5.4.6 Stress

Typically, one would expect to see the strongest correlates of stress on vowels. Empirically, studies have shown that it is the quality, pitch, amplitude, and duration of vowels that changes when they occur in stressed versus unstressed position. Thus, stress effects should not be found on consonants (Hardcastle & Marchal, 2012). None of the NAE vowels showed effects of syllable stress that would have influenced M.S.’s

production accuracy. Consequently, it may seem odd at first glance that postvocalic /ɪ/ would be affected by syllable stress. However, the NAE vowel /ɜ-/ and its unstressed counterpart /ə-/ are rhotacized, meaning they are stressed and unstressed variants of schwa that have a postvocalic /ɪ/ quality to it. Thus, the articulation of an /ɪ/ sound is part of the vowel (acoustically, /ɜ-/ has a lower F3, which is associated with /ɪ/). So it would make sense that the production of /ɪ/, being partly vocalic, would be affected by stress. The results revealed that in unstressed position, M.S. was much more likely to vocalize /ɪ/. In the token *burger*, for example, the first syllable that bears stress *bur-* would be on target, whereas the second syllable *-ger* was vocalized, thus the output was [bɜːgɜ]. M.S. may be driven to vocalize his postvocalic /ɪ/s more if the preceding vowel is unstressed, and therefore not salient enough, like in the words *banker*, *beggar*, or *radar* (Celce-Murcia et al., 1996).

5.4.7 Self-Assessment and interview: extra-linguistic factors

In addition to the intra-linguistic factors discussed above that were investigated to find out how and if they influenced M.S.'s production for NAE segments, the 30 minute interview that was conducted after the production study was utilized to learn about cognitive and psychological factors that have a bearing on M.S.'s production performance. During the interview, M.S. gave some information on his background of learning English, and psychological variables of L2 pronunciation learning, such as motivation, attitude, aptitude, identity, personality, and anxiety. Furthermore, the interview contained a self-assessment on his pronunciation ability in which he shared how he perceived his own accent in comparison to that of native speakers of English with specific focus on some particular sounds. This information provided insights about the

constitution and state of M.S.'s interlanguage grammar (section 5.5) beyond the analysis of accuracy scores of certain NAE segments.

Language Learning Background

In the questionnaire and the interview M.S. revealed information about factors pertaining to Age of Onset (AoO), Length of Residence (LoR), Age of Instruction (AoI):

“I spent a year in Australia, and I spent three years in the US, going to university, and then moved to Winnipeg in 1996, but before that I only learned English in high school, and you know, the teachers back in the day: they couldn't care less about pronunciation. I'm sure they were convinced none of us were ever going to go abroad anyway or use English much after graduation, so there was just no need for that.”

M.S. has spent over 25 years in English speaking countries all over the world, ranging from Australia to the USA and Canada where he has resided since 1996, so he was ~25 years old when he first received input from native speakers of English from two different continents. M.S. did not receive any formal pronunciation training in high school and he was only ever exposed to German speaking teachers throughout German high school, where he started learning English at the age of 11.

Psychological factors: Attitude, Identity, Aptitude, Motivation, Personality, Anxiety

On Identity:

Moreover, the interview revealed details about M.S.'s cultural identity, attitude, and motivation as a L2 learner of English: M.S. still holds German citizenship and expressed a very tight connection to his relatives in Germany and German culture:

“I mean, this is where I grew up and I think it’s important to keep your roots, your heritage. And that’s what I like to do. I like living in Canada, but the good thing about this country is that no one forces you to assimilate to this culture, you can bring yours and you’re welcome to do that.” It was evident from the interview that he identifies as German, despite having spent half of his life in Canada, and the strong attachment to his family supported this.

On Aptitude (Talent) and Motivation:

M.S. further explained that he did not really want to sound like a native speaker of English and that he never attempted to do so: “I must say I’ve never been one of those extremely talented people that watch movies and pick up accents from other people they interact with. I mean, I’m sure I don’t have that sort of talent. And I don’t actually care too much about it. The most important thing to me is that people understand me and I have never had bad experiences because of my accent, so I don’t see the need to change it.”

On Personality and Anxiety:

M.S. was also asked how performing in NAE as an L2 made him feel and whether he interacts frequently with people: “Well, it’s of course difficult when you don’t know how to express yourself as well as in your first language, so I’m definitely not at ease with English in every situation, but usually that doesn’t keep me from speaking. I think there are always ways to describe what you want to say, you don’t need a lot of vocabulary, even though I have enough of that, but I don’t find it difficult to talk to people, even though I would say the majority of my contacts is probably in Germany and elsewhere over the world. In Victoria, of course you speak a lot of English, but I can’t say that I’m

the kind of guy who seeks out conversations and hangs out with a lot of people that would require me to speak it all the time.”

The table below summarizes the extra-linguistic factors pertaining to overall pronunciation M.S. reported throughout the interview and language learning background questionnaire.

Table 5.6 Extra-linguistic factors affecting M.S.’s L2 speech

M.S. extra-linguistic factors influencing L2 speech production	language learning background			psychological variables					
	LoR	OoA	AoI	attitude	aptitude	motivation	identity	personality	anxiety
√ or -	-	-	√	√	√	√	√	-	-

Self-assessment of Problematic NAE Segments (M.S.’s pronunciation perception):

Furthermore, the self-assessment served as a means to gain more insight into features of M.S.’s pronunciation on the micro level, as he reported strengths and weaknesses of his L2 pronunciation ability. In Chapter 2.3 it was explained that both phonetic as well as phonological accuracy errors create a foreign accent (Archibald & Libben 1995), except that the L2 learner “...cannot monitor [phonetic] errors” (Archibald & Libben, 1995: p. 239), because he or she typically does not hear them, which is in line with what M.S.’s interview showed.

As Derwing (2003) found, M.S. was able to comment on a number of segments of NAE that pose difficulties for him, albeit that he did not identify all of them. These segments were mostly consonants about which I learned some rather specific things that

the production data alone could not have accounted for. M.S. was asked whether there were segments that he to this day finds very difficult to pronounce, to which he said: “I find the difference between /w/ and /v/ difficult and it was not until I moved to Vancouver and a friend pointed out to me that /w/ and /v/ are not the same in English... I guess now I try to make that distinction, but I also find it difficult to always think of the difference.” He also commented on his production of NAE interdental: “The <th>-sound... that was the first sound that our teachers always drew our attention to in school, but for some reason I always associated <th> sound appearing more in the beginning of words than at the end of words. I know this may not be the case, but actually, I think that if I put enough effort to make my <th>s sounds right in the beginning of a word I don’t have to put as much effort into making them sound right at the end of words”. Moreover, he also commented on NAE segment /ɹ/, which he never produced as German /ʁ/, because they are “very different”, he found. “English /ɹ/ is just very difficult to make for me, I can’t bring my tongue to make this sound, but I know it’s not like my German <ʁ>. But I notice that whenever there’s an <ʁ> sound around it can be very difficult for me to pronounce, I try not to say it too often [laughs], because, yeah, even though that is hard, because English has so many of them”.

Furthermore, he also indicated that he was certain that this would not change in the future: “I know that if I really put a lot of effort I can pronounce most of the English sounds, but there are just these few that I know... that’s just never going to change no matter how hard I try. Also, it is very difficult to focus on the words you’re trying to say, form a grammatical sentence, and also make sure that the pronunciation is like that of a native-speaker. I mean I can’t focus on all of these things at once.” Here, M.S. revealed

some details about how some segments may have reached a stage of fossilization so that they will no further change (Selinker, 1972).

Other than these segments he did not comment on anything that he found difficult to produce. He mentioned: “There is probably a lot of things I can’t even hear, but that also does not matter to me then. If I can’t hear it, it’s probably not that problematic if I don’t say it correctly, and I have friends who would have pointed that out to me, I think”.

Overall, from the qualitative comments on his segmental production, one can say that he identified some of the “classics” that have been primarily focused on in previous studies involving L2 pronunciation of NAE by native speakers of SHG. It was striking that all of the segments M.S. commented on were consonants – he did not find any NAE vowels difficult to pronounce, at least he mentioned that he had not picked up on any. This could be explained because of the different nature vowels and consonants have, of course. Since consonants allow significantly less room for production variety, as opposed to vowels (Chapter 3.4), he seemed to perceive and assess his vowel productions to be target-like and good exemplars of NAE that he did not experience any particular difficulties with.

The table below summarizes the segments M.S. identified as problematic in the interview and classifies them by interlanguage phenomena (as described by Tarone in Chapter 2); and will be further discussed in the following Section 5.5 on Interlanguage:

Table 5.7 Self-Assessment of Problematic NAE Segments

FOSSILIZATION of NAE segments			
Segments M.S. identified as problematic in the interview:	/ɪ/	/θ/ and /ð/; /ɪ/	/w/-/v/ contrast
Interlanguage phenomenon:	Approximation → v	Avoidance → /θ/ and /ð/ in coda position ; /ɪ/ “difficult to make”	Overgeneralization

Overall, the most substantial extra-linguistic factors that M.S. reported to make up his L2 speech productions were psychological factors such as motivation, identity, instruction, anxiety, personality, and aptitude, all of which cannot be quantified, but provide further insight into factors that can possibly account for his L2 speech production and the maintenance of a foreign accent. Combined with these, M.S.’s assessment of some of his own pronunciation difficulty of some segments and those that escaped his attention (e.g. he did not mention anything about vowels) loosely followed the patterns observed and reported in the segmental analysis, and can be linked to interlanguage phenomena such as approximation, avoidance, and overgeneralization (Tarone, 1987) of certain NAE segments like /ɪ/, /θ/ and /ð/, and /w/.

5.5 M.S.’ s Interlanguage Grammar

Considering M.S.’s accuracy scores for the NAE segments and the substitution types he prefers for each segment, a picture of how these NAE sounds are represented in his phonological mind starts to emerge. These representations are also highlighted with the interview and self-assessment M.S. gave on his pronunciation.

Two of the research questions that were raised in Chapter 1 pertained to how M.S.'s production of NAE segments can provide insights into his interlanguage grammar, the structure thereof, and whether his interlanguage grammar is 'variable' and therefore dependent on task type. What emerged from the analysis of the results were indeed conclusive insights into M.S.'s *interlanguage grammar*. In this section on IL, I will make two important points about his IL grammar; one pertains to the hypothesis that production errors can be accounted for by L1 transfer exclusively, something which is easily refutable, and the second point pertains to the idea of variable competence (Ellis, 1984 ; Tarone, 1987 ; Major, 2001) vs. variable performance (Gregg, 1990 ; Chomsky 1965).

In order to understand what the state of M.S.'s interlanguage grammar looks like, we have to consider that M.S.'s productions of NAE segments revealed that his phonological representation for NAE sounds is comprised of segments that are part of his L1, segments that are part of his L2, and segments that are located in between the L1 and L2, sometimes being a mix of the two. The figure below illustrates this composition:

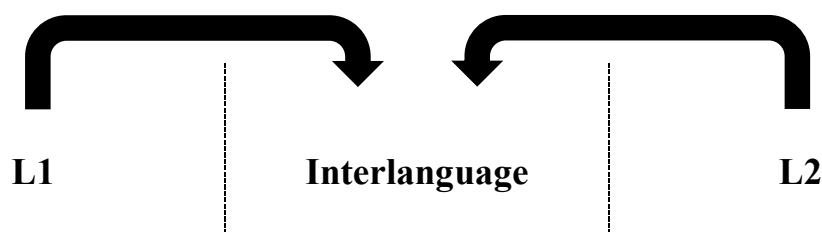


Figure 5.1 Influence of the L1 and L2 on the Interlanguage Grammar

The table below shows the NAE sounds that were of interest to this study and M.S.'s preferred ways of producing them, which provides insight into how these sounds are stored in his interlanguage grammar. The production of the 3 NAE segments in the

left column (L1-like) suggested that M.S. realized them as L1 sounds. In the middle column are the segments for which M.S. showed variation in production. The brackets [...] indicate the output he produced for the target segments on the left in slash brackets [...]. The column on the right shows the segment of NAE that M.S. has successfully acquired and which is nativelike in production:

Table 5.8 The State of Segments in M.S.'s Interlanguage Grammar

Constitution of M.S.'s Interlanguage Grammar for NAE Sounds

L1-like	IL	L2-like
/u:/ → co-articulatory effects	/æ/ - [æ ~ æ̃ ~ ε] → free variation	/v/
final devoicing of coda obstruents → only on target if following onset is voiced	/ʌ/ - [ʌ ~ a] → free variation	
epenthesis and non-reduction of [t]	/p,t,k/ - [p,t,k ~ p ^h ,t ^h ,k ^h ~ p̄,t̄,k̄] → free variation	
	/θ,ð/ - [θ ~ s ; ð ~ z] → conditioned by syllable context	
	/ʃ/ - [ʃ ~ l] → conditioned by syllable context	
	/w/ - [w ~ v ~ v̄] → free variation; OR vowel environment	
	/ɪ/ - /ɪ/ - [ɪ ~ v ~ ∅] → free variation	
	/ɪ/-vocalization [ɪ ~ e]	

First of all, looking at the table, and after careful analysis of the results, we can see that M.S. is not exclusively transferring L1 SHG sounds to his L2 NAE. We have known since the early criticism of *Contrastive Analysis* that we cannot simply compare the sound system of an L1 and an L2 and predict that the only errors will occur due to transfer (Archibald & Libben, 1995), and unsurprisingly, this is the case for what was observed in M.S.'s speech production across three different task types. As Tarone (1987)

argues, M.S. has a large amount of NAE segments that are neither exclusively part of his L1 or his L2. Rather he is approximating, overgeneralizing, or avoiding certain L2 sounds and sound patterns. To give a few examples from the table above, we find that M.S. overgeneralizes the occurrence of /w/ vs. /v/, often having them in free variation where they are contrastive in the L2, and he approximates /w/ with a labio-dental approximant [v]. The same observations can be made for /æ/ for which he uses [æ ~ æ̣ ~ ε] in free variation, and /ʌ/ where [ʌ ~ a] also exist in free variation. The same applies to postvocalic /ɪ/ which is realized as [ɪ ~ e] in his production, the interdentalals /θ/ and /ð/, represented as [θ ~ s ; ð ~ z], and NAE /t/ being produced interchangeably as [t ~ l] across tasks. Moreover, NAE /p,t,k/ in coda position is produced as [p,t,k ~ p^h,t^h,k^h ~ p̄,t̄,k̄]. As can be seen, some of these NAE sounds that exist in free variation to him do not form phonemic categories, where English would contrast these sounds: /æ/ and /ε/, /w/ and /v/, /θ, ð/ and /s, z/, and /ɪ/ and [v~w]. As explained in the previous section, these sounds are contrastive phonemically, but to M.S. they represent the same category, because he has never established separate categories for these NAE sounds. The patterns of free variation and approximation of these phonemically contrastive sounds of NAE let us draw conclusions about how they are represented in his phonological mind, where they do not contrast. This representation is part of M.S.'s interlanguage grammar. This will be further explained in the section on the Speech Learning Model below.

Three segments of NAE that M.S. produced indicated that they were mostly L1 like, as M.S. consistently produced the same sounds for them: his NAE /u/ segments were productions of the SHG [u] sounds with very few on target productions, his devoicing of underlying voiced obstruents in coda position was consistent throughout tasks (more on

co-articulation below in “No Variable Competence”), and his speech productions exhibited [t] epenthesis and non-reduction of /t/ consistently through tasks with the exception of one token (NAE *actually*). The only segment of NAE that was considered “acquired” by means of accuracy threshold was the vowel /ɒ/, which is absent from SHG and for which he seemed to have a relatively stable category established.

Concluding from this table, we can see that M.S.’s interlanguage grammar, the mental representation of NAE sounds, is comprised of L1 segments, L2 segments and segments that are part of the interlanguage, and therefore in between the L1 and L2. This revealed the state of his interlanguage grammar after prolonged exposure to English in an English-speaking environment for over 25 years. We can see that nativelike pronunciation was only achieved for one segment of NAE, whereas for the rest of NAE segments (11), M.S. reached a certain stage in attainment: 3 segments (L1-like) remain purely L1-like, meaning that M.S. has never acquired the NAE sound structures for them, whereas for the other 8 (interlanguage), it can be observed that at one point he was in the process of acquisition, which led to approximations and overgeneralizations, until he no longer progressed and a final stage was reached. In other words: his phonological grammar acquisition plateaued and became fossilized (Selinker, 1972), his interlanguage grammar mostly consisting of approximations and overgeneralizations that are neither purely L1, nor L2-like, as they exhibit structures from both the L1 and L2 (overgeneralization, e.g. $\theta \sim s$; $\delta \sim z$) or neither the L1 and the L2 (approximation, e.g. [v] \sim /w/). This is indicative of a pattern in his L2 speech production that exists and will continue to persist.

No ‘variable competence’

Now that the accuracy scores, M.S.’s preferred substitutions, and his self-assessment on NAE pronunciation attainment have painted a picture of the constitution of his interlanguage grammar, we turn to the discussion of whether M.S. has one stable interlanguage grammar with variable performance, or if his linguistic competence of NAE sounds is variable. The ‘variable competence model’ (Tarone, 1987 ; Ellis, 1984) suggests that the L2 learner does not use a single homogenous interlanguage grammar, but, in fact, several, depending on the discourse situation and task formality. Within this model, M.S. would not have a fixed homogenous system of his NAE interlanguage; rather, new forms of NAE sounds would keep entering this system and co-exist with an old form. This suggests that M.S. can have several representations of NAE sounds (several grammars), which can be altered, become activated depending on the discourse situation (task formality), and finally surface in his speech production. To illustrate this, I will use the example of the English interdentalals: the ‘variable competence model’ would claim that in a formal task, M.S would be able to produce on-target L2-like interdentalals NAE /θ/ and /ð/ in a nativelike manner, activating a grammar in which these sounds exist, but in a spontaneous production task, M.S. would substitute the interdentalals with [s] and [z], respectively, and become more off-target-like due to activation of a different grammar:

This hypothesis that L2 learners have ‘variable competence’ (Tarone, 1987; Ellis (1984), and thus employ several grammars depending on discourse type and task formality, was tested against Gregg’s hypothesis (1990), who draws on the Chomskyan notion that production variability (performance) does not reflect general acquired

competence of the L2. Thus, as explained in Chapter 2 (section 2.3) we have two opposing views of the constitution of a learner's interlanguage grammar.

As this is central to the research questions of my thesis, in order to draw conclusions about the state of M.S.'s interlanguage grammar, the obtained production data necessitated careful analysis to differentiate *performance* from *competence*. The analysis demonstrated that we have to reject the notion of 'variable competence' (Tarone, 1987 ; Ellis 1984) and suggested that M.S. only has **one** underlying competence (Gregg, 1990 ; Chomsky, 1965). This claim will be supported and evidenced in the following paragraphs.

Two major observations stood out which lead me to side with Gregg on the notion that M.S.'s interlanguage is a single, fixed, homogenous system. The first reason that would refute the notion of variable competence is the sheer fact that no significant differences across task formality could be observed for the majority of segments. The authors in favour of the 'variable competence model' have argued that, depending on formality of discourse, task type, and social context, the interlanguage grammar of L2 learners differs. Some studies (Gatbonton, 1975 ; Dickerson, 1974) suggested that the most formal task would show the highest accuracy scores, whereas others (Silveira, 2007 ; Major, 1994 ; Lin, 2001) suggest that the least formal task shows the highest accuracy scores, but no matter which holds true, it is assumed that there is underlying variable competence which is related to the task type formality. At first glance, M.S.'s productions of the voicing contrast in coda position seemed to significantly improve from the most formal task (wordlist) to the least formal task (semi-spontaneous production), for which he achieved accuracy scores of 0.8% in the wordlist task, to 27% in the

sentence reading task, to 40% in the semi-spontaneous production task. The other segment that exhibited apparent task type effects was epenthetic [t]/non-reduction of /t/, which was consistently at 0% in the wordlist and sentence task, but improved dramatically in the spontaneous speech (76%). Irrespective of these two NAE segments, no task type effects could be observed for the other segments (except minor ones for /æ/ that could be mostly attributed to other intra-linguistic factors; see section 4.1). However, observing apparent task type effects for these two segments, which Tarone (1987) and Ellis (1984) would ascribe to ‘variable competence’, it was not *competence* (linguistic knowledge) that changed throughout the different tasks, but rather M.S.’s *performance*.

Taking a closer look at the two examples of 1. final devoicing of coda obstruents and 2. epenthetic [t]/non-reduction of /t/ that showed task type effects, it surfaces why M.S. does not have variable competence or access to different grammars in different tasks. Instead, differences in production performance could be attributed to external factors (intra-linguistic factors) conditioning the improvement of his production across tasks types:

1. M.S.’s accuracy score for the on target voiced obstruents in coda position was at 1% in the wordlist task, but improved dramatically in the sentences and semi-spontaneous task, where it reached its peak (40%). Thus, in the wordlist, he only produced one instance of coda obstruent voicing. In the sentences, this seemed to have improved, as his performance accuracy reaches its peak. When taking a closer look, one can see that $\frac{3}{4}$ of all the voiced codas he produces in the sentence task can be attributed to co-articulatory effects of following voiced consonants, or possibly resyllabification (see section 4.5) with following vowels, e.g. “*She had bad **an**xⁱety.*” The initial vowel of the

term *anxiety* influences the production of the coda in *bad*, something which – for comparison – he did not get right in the wordlist task. From the wordlist task, in which he produced only a single voiced coda obstruent, it can be concluded that M.S. does not have the English coda voicing contrast stored as linguistic knowledge (competence) of English. Even though for most of the tokens he was given there was a grapheme-phoneme correspondence, like in <*bad*>, which is pronounced as [bæd] by a native speaker of English, he would consistently devoice underlying voiced coda obstruents. This serves as evidence to demonstrate that it is purely by effects of the surrounding phonetic environment in which these coda obstruents appear that M.S.'s performance improves; a result of re-syllabification. M.S.'s grammatical representation for the NAE voicing contrast, however, is that of his L1 German, which only has voiceless or devoiced coda obstruents.

2. For epenthetic [t], M.S. accuracy score was 0% in the wordlist and 0% in the sentence task, but then seemed to improve dramatically in the semi-spontaneous production task (76%). This seems significant and might suggest at first that in spontaneous production his accuracy is higher because M.S. activates different grammatical knowledge, thus the knowledge that English reduces /t/ in clusters and does not exhibit any /t/ in words ending in <-tion>, so M.S. does not epenthesize [t]. However, I found that M.S.'s accuracy scores improves so dramatically because he uses the NAE word *actually* repeatedly, for which he successfully reduced the /t/. This one and only token can hardly be representative of his competence for this segment, though. For all other instances throughout all three productions tasks, M.S. consistently epenthesized [t] or failed to reduce the cluster, suggesting that he never really 'acquired' this reduction

pattern and that the token *actually* did not seem to pose any difficulties to him and seemed to be the result of a frequency effect: many native speakers of NAE will use this word frequently, and so will he, as it is a discourse marker.

While these are only a few examples from the data, Tarone (1987) and Ellis (1984) argue that variable production accuracy across different task formalities is indicative of an individual employing different *grammars*, thus, activating different grammatical competences in a given discourse situation. Since we seemingly encounter task type effects in M.S.'s production of the voicing contrast in coda position and for epenthetic [t]/non-reduction of [t], Ellis and Tarone would argue for variable competence. They would claim that M.S.'s performance improving in the sentence task for the English coda voicing contrast can be attributed to him activating a different grammar that increases his accuracy score. The same would apply to M.S.'s seemingly dramatic improvement from the wordlist and sentence task to the semi-spontaneous production task for reduction of [t]. However, when we take into account that his production accuracy for both of these segments improves from 0% and 1%, respectively, to 76% and 40%, respectively from the most formal to the least formal task, we have to consider that this is not due to M.S. activating different grammars in different task contexts. Rather, we have external factors (intra-linguistic) that influence his production, but underlying, the representation of these sounds in his phonological mind is consistent and stable. Because we can see apparent task type effects for the two segments discussed above, we can argue that his *performance* certainly varies, but this cannot be attributed to 'variable competence'. The fact that M.S.'s production accuracy for the voicing contrast improves for voiced coda obstruents in the sentence reading task is an effect of the phonetic

environment, as an adjacent word with a voiced onset or a vowel would trigger him to resyllabify, thus, moving the coda of NAE *bad* to the onset of NAE *anxiety*. making it look like he does have coda voicing. But it is only with the effects of phonetic environment influencing his coda production, and as we can tell, in isolation, none of his productions demonstrated that M.S. has a grammatical representation for coda voicing contrast, in which he contrasts /t/ from /d/ or /b/ from /p/ in syllable-final position, as the auditory evidence suggested. This means that we can say something about his competence, his linguistic knowledge of this phonological process of English, and that his production of this voicing contrast improves in the sentence task, but only due to external factors conditioning this production improvement. M.S.'s representation of this process does not change and is therefore not variable. This is why I argue that such task type effects can be attributed to external factors (intra-linguistic) that can result in variable performance, but task type effects are not an argument for 'variable competence'.

In fact, if we wanted to argue for 'variable competence' (Tarone, 1987 ; Ellis, 1984), we would have to ignore the intra-linguistic factors that were analyzed in the result section, some of which were shown to have influenced M.S.'s production of the NAE sounds of interest examined in the study. Undoubtedly, we have a variety of intra-linguistic factors like orthography, vowel environment or syllable context affecting M.S.'s production, which results in differing accuracy scores across task types (for the two segments mentioned above), but looking more closely, we can say that it is these very factors that can only account for 'variable **performance**'. Similarly to native speakers, production errors stemming from the influence of intra-linguistic factors can

account for variable *performance*, e.g.: if a native speaker aspirated their voiceless stops in the wordlist task like in *pet* [pet^h] or *back* [bæk^h], but would not do so in connected speech, we can assume that this is not due them having several competences of how to produce stops, but because they hyperarticulate words in isolation causing unreleased stops to become released (Blevins, 2004). In connected speech, native speakers of English would likely not hyperarticulate then, so their *performance* can differ across different task formalities, but there is no reason to believe that it is their linguistic knowledge of producing stops that fluctuates. As with M.S., their speech production is influenced by external factors like orthography, phonetic environment, stress, or frequency, etc. These errors do not stem from his interlanguage phonological grammar, and thus his competence (Chomsky, 1965). In other words, notwithstanding task type formality influencing performance, we can assume that the mental representation M.S. has of these NAE sounds does not change from task type to task type, the underlying competence remains the same, and there was no evidence borne out in the results that suggested otherwise.

Based on these arguments, I concluded that M.S.'s has one fixed, homogenous IL grammar of NAE, and Table 5.8 (from the beginning of the section) gives us an idea about the constitution of the grammar phonologically. We see a fixed system of an interlanguage grammar that has plateaued and is fossilized. For some segments, like epenthetic [t]/non-reduction of [t], and voicing contrast in coda obstruents, as well as for the production of the vowel /u:/ M.S. exhibits L1 transfer; whereas for some others we can observe patterns of overgeneralization like /æ/ and /ɛ/, which exist in free variation. Furthermore, we observe patterns of approximation, such as when he attempts accurate

/w/ and /ɪ/ productions by producing an in-between sound [v] that is neither native to his L1 nor his L2; and how he produces other sounds in free variation, like /w/ and /v/, or /s/ and /θ/, all indicative of him having established a systematic **IL grammar** which is not malleable by task type formality, ergo, context or discourse. And lastly we have one segment that he has acquired which is truly **L2 (native)-like**, which is NAE /b/. The constitution of a homogenous interlanguage can be further explained with regard to the extra-linguistic information obtained from the subject himself. There is little reason to believe that at one point M.S. will achieve nativelike performance or perform nativelike in a certain task type and return to free variation again in another. The NAE sounds that constitute his interlanguage grammar are part of the linguistic knowledge he has acquired of his L2, which can be regarded as similar to a fixed, homogenous competence he has in his L1. This system may differ in form from that of a native speaker of NAE, but it is just as functional, rule-governed, and systematic.

5.6 Speech Learning Model

In this section, we will now turn to the predictions made by the Speech Learning Model on the ease/difficulty of acquiring *similar* and *new* L2 sounds, which will reveal that there are a variety of factors at play that make it difficult to predict which L2 sounds will be easy or difficult to acquire. While Flege's model relies on phonetic similarity between an L1 and an L2 sound and the formation of L2 phonetic categories, we will learn that the role of perception is crucial; a role that Flege cannot account for in his model.

This section evaluates the predictions made by Flege's Speech Learning Model on the individual segments of NAE that were under investigation in this study. Following chapter 2.3.1 and 2.4.9 on Flege's Speech Learning Model and its predictions, *similarity* or *newness* of an L1 sound and an L2 sound was defined based on the comparison of the phonetic distance between the NAE sounds of interest and the closest SHG sound (usually the favoured substitution described in the literature review, e.g. SHG /v/ for NAE /w/). The SLM predicts that *new* sounds of NAE will be *easy* for M.S., and therefore more accurate in production. In contrast, *similar* sounds will be *difficult* to produce due to perceived similarity between them and existing SHG sounds (equivalence classification).

As established, the NAE sounds that differed from SHG sounds in very few acoustic properties happened to be the ones that have a phonemic equivalent in SHG, which is why they were labelled as *similar* and therefore hypothesized to be difficult to produce (/u:/, [ɬ], and [p̄, t̄, k̄] in coda). From M.S.'s production data, it emerged that the NAE vowels /ɒ/, /ʌ/ were of interest as well, so they were added to the table of section 2.4.5 on the Speech Learning Model after being measured and analyzed. Given the phonetic similarity of NAE /ʌ/ to SHG /ə/, which can be considered an unstressed counterpart, it was established that these sounds are phonetically *similar* and therefore should be *difficult* to acquire. Comparing NAE /ɒ/ to its closest SHG counterpart, it was found that it differs only slightly from SHG /ɔ/. support of this, one can employ the example of the NAE vowels /ɔ/, /ɒ/, and /ʌ/. Because these phones exhibit a high degree of phonetic similarity, they tend to be merged in some parts of North America (Labov, Ash, & Boberg, 2006) and therefore do not represent distinct phonological categories. By

analogy, it could be concluded that the closest SHG sound /ɔ/ would be *similar* to NAE /ɒ/ and therefore *difficult* to acquire. On the other hand, the *new* sounds of NAE (/æ/, /θ/ - /ð/, /w/, and /ɹ/) are the ones that have distinct phonological categories in NAE, but not in SHG.

In total, out of 9 NAE segments, 5 were predicted to be difficult for M.S., and 4 were predicted to be easy. All segments that were reported in the result section are brought together in the table below and the hypotheses on predicted vs. attested ease of acquisition based on the SLM are presented:

Table 5.9 Accuracy of SLM Predictions

NAE sound	/æ/	/ʌ/	/ɒ/	/u:/	[p̄, t̄, k̄]	/θ/ - /ð/	[θ]	/w/	/ɹ/
closest SHG sound	/ɛ/	/ə/	/ɔ/	/u:/	[p ^h , t ^h , k ^h]	/s/ - /z/	/l/	/v/	/ɻ/
New? Similar?	new	similar	similar	similar	similar	new	similar	new	new
Predicted ease of acquisition	easy	difficult	difficult	difficult	difficult	easy	difficult	easy	easy
Acquired?	X	X	√	X	X	X	X	X	X
Prediction confirmed?	X	* (√)	X	√	(√)	(X)	√	X	X

* (√) or (X) in parentheses indicated that the predictions were confirmed or not confirmed according to accuracy threshold, but also reflect that this could not be determined with high confidence.

In the ‘Prediction confirmed?’ row, the check marks indicate that the prediction was confirmed, e.g., NAE /u:/ is *similar*, and therefore was difficult to acquire. The check

marks and Xs in parentheses mean that the prediction was confirmed or not confirmed but there were other circumstances which conditioned this: e.g., NAE /ʌ/ is *similar* and was therefore predicted to be *difficult*. According to accuracy threshold /ʌ/ had not been acquired, but the actual accuracy score of 65% left room for interpretation as to whether or not a new phonetic category has been established. Similarly, NAE /θ/ - /ð/, which are *new* and therefore predicted to be *easy*, were not acquired according to accuracy threshold, and so the SLM prediction was not confirmed, but the interdentalals were produced with such a high accuracy in onset position that they can be considered phonetically acquired, because syllable context seemed to condition M.S.'s production. The Xs mean that the prediction was not confirmed, e.g., /ɪ/ was predicted to be *easy*, but it did not reach the accuracy threshold.

It is important to make a distinction between mere production and the establishment of a new phonetic category. The SLM makes predictions on the establishment of a new phonetic category, and therefore on acquisition, not on the simple production of a speech sound. NAE sounds which are *new* were predicted to be easy to establish a new phonetic category for. NAE sounds which are *similar* were predicted to be difficult to establish a new category for.

On the opposite end, /w/'s accuracy score suggested that M.S. was merely able to produce a few instances of [w] (17% of productions total), but the high rate of substitutions with [v] and [ʋ] suggested that M.S. seems to equate this sound to the familiar L1 category of /v/, which is why the formation of a /w/ category in his L2 English seems to be blocked (Flege, 2003).

Similar sounds [u], [ʊ], [ʌ], [ɨ], and [p̄, t̄, k̄] were predicted to be difficult due to phonetic similarities to SHG counterparts. Two of the predictions for these sounds were confirmed ([u] and [ɨ]), two were partly confirmed ([ʌ] and [p̄, t̄, k̄]), and one was not confirmed ([ʊ]). For the two predictions that were confirmed, M.S. can be said to have not established a new phonetic category because he perceives NAE /u/ and NAE [ɨ] to be an equivalent of the respective L1 categories SHG /u/ and SHG /ɨ/. On-target productions of these sounds, in addition to being nowhere near threshold (at 40% and 24% accuracy respectively), were conditioned by syllabic context. NAE [u] was only produced on-target when following a palatal or alveolar consonant as an effect of co-articulation⁸, and NAE [ɨ] was produced on-target in coda position only. For the two predictions that were partly confirmed ([ʌ] and [p̄, t̄, k̄]), accuracy scores reflect the difficulty of acquiring them, given that they were not at threshold (65% and 69%, respectively), but because their accuracy was so near threshold, it is possible that at some point M.S. was on the way to category formation.

Additionally, it is important to note that these sounds, which were predicted to be difficult, had a higher accuracy rate than some sounds which were predicted to be easy; indeed, these difficult sounds have the second-highest rates of accuracy of all M.S.'s segments. Finally, /ʊ/, being a *similar* vowel to SHG /ɔ/, was never substituted. This is counter to the prediction made by the SLM that this *similar* sound would be difficult to acquire because it is completely absent from SHG. /ʊ/ was, in fact, the only NAE sound

⁸ Co-articulatory effects conditioned the accuracy of M.S.'s NAE [u] productions. Thus, when measured in isolation and not taking phonetic environment into account, the acoustic analysis revealed that M.S.'s productions were target-like. However, one has consider here that strictly speaking he was never on-target, as the native speaker produced [u]s following palatals and alveolars that were even further fronted than M.S.'s.

M.S. produced with an accuracy score above the threshold (92%). Here we have to question what M.S.'s perception of this sound may be. Despite the phonetic similarity, perhaps there is a greater perceptual dissimilarity that he perceives which could account for his accuracy in producing this sound.

Concluding, an interesting trend could be observed in that the sounds which were predicted to be difficult due to similarity had three of the four with the highest accuracy scores: /b/ at 92%, [p̄, t̄, k̄] at 69%, and /ʌ/ at 65%. In contrast, none of the *new* sounds which had been predicted to be easy, had accuracy scores indicating that a phonetic L2 category had been established. Especially NAE /w/ stood out, because SHG has no phonetic near-equivalent to this sound, but the accuracy score was only 17%, the lowest of all segments, as a result of the fact that M.S. seemed to totally equate this to his L1 /v/. The fact that *new* sounds did not make it to the accuracy threshold can be accounted for by two points: 1. that M.S. did not establish new phonetic categories for these sounds and 2. that they are phonemic, and therefore contrastive, in the L2. What would otherwise be phonetic errors (such as with the *similar* sounds) become phonemic errors, and this is something Flege's model does not account for. An example would be M.S.'s production of NAE /ɪ/, which he does not substitute with the closest SHG equivalent /ɪ/, but with [v]. In turn, he M.S. also substitutes NAE /w/ with [v], making it a phonemic error. The below diagram shows the conflicts that M.S.'s productions of *new* sounds have with multiple NAE phonological categories. The top and bottom rows are different NAE phonological categories, and in between are M.S.'s productions. Lines indicate what phonemes the productions can be linked to:

NAE phonological category 1:	/æ/	/θ/	/ɪ/	/w/
M.S.'s substitution:	[ɛ]	[s]	[v]	[v]
NAE phonological category 2:	/ɛ/	/s/	/w/	/v/

Establishing new phonetic categories for sounds that are phonemic in the L2 is key, because although it does not matter if NAE /l/ is produced as [l] or as [ɫ], it does matter if NAE /θ/ is produced as [s] and not [θ]. Flege's model does not consider this phonological aspect of L2 sound acquisition.

Finally, the role of perception needs to be further explored because the *new* sounds are the most phonetically dissimilar sounds and Flege assumes that accurate perception precedes accurate production, therefore tying in phonetic similarity with perceptual similarity (Flege, 1995 ; 2003). This was not confirmed in the findings of the thesis, though specific details about M.S.'s perception are unknown. Flege acknowledges that his SLM cannot account for the interplay between phonetic and perceived (dis)similarity, which he has criticized about his own model (Flege, 2016). M.S.'s production data show that this incomplete aspect of the SLM is worth exploring and developing further. Furthermore, it could be established that the SLM did not accurately predict the ease or difficulty of acquiring *new* or *similar* sounds. Especially for those NAE sounds that were predicted to be *easy* to acquire because they were more

phonetically distant from M.S.'s L1 SHG categories, none of the predictions were confirmed in the findings. In fact, not a single NAE sound that had been predicted to be *easy* could be measured as “phonetically acquired”. Therefore, we must conclude that *phonetic* similarity or dissimilarity between an L1 and an L2 sound alone cannot predict acquisition difficulty. *Perceived* similarity does not seem to equal *phonetic* similarity as M.S.'s production data clearly shows.

5.7 Global Accentedness, Comprehensibility, and Intelligibility

The native speaker judgment ratings served to rate M.S.'s L2 speech production for accentedness, comprehensibility, and intelligibility. Despite ratings indicating that there was a ‘noticeable foreign accent’, it is important to point out that M.S. achieved high comprehensibility and intelligibility ratings, which should be the only goal in L2 English attainment. Not only has it been established that nativelike attainment is hardly ever possible (Munro & Derwing, 1999 ; Hyltenstam & Abrahamsson, 2003), but that there is simply no need to attempt accent reduction as long as one can effectively communicate in the L2, at least in the English context. Bringing together the native speaker judgments with the content of M.S.'s interview, one can conclude that M.S.'s goal to be intelligible and comprehensible has been fully met, despite maintaining a foreign accent. Achieving high comprehensibility and intelligibility scores means that errors do not hinder communication. We know from M.S.'s IL grammar that he seems to perceive phonemic contrasts and attempts accurate production thereof, which cannot be said for phonetically similar sounds like most vowels, [ɪ], and [p̄, t̄, k̄], for example. In the interview he does

not mention that these sounds are problematic for him and his production data shows that he may not be aware of the acoustic dissimilarity of, say NAE /u:/ and SHG /u:/. This might very well be because a) it simply does not matter, b) M.S. seems to know that producing these in a SHG like manner will not hinder his comprehensibility and intelligibility and c) the preservation of the phonetic quality of these segments, which is that of SHG, may serve as an identity marker to him, which allows him to stay attached to his L1.

While it is unlikely that he ignores acoustic differences between NAE and SHG sounds deliberately, one can assume based on his statements he has shared on lack of motivation to sound more nativelike in his L2, that he does not make a conscious effort to attempt to perceive NAE phonetic details either. His goal to be understood is fully met, and he preserves some of the phonetics of SHG L1 sounds, in line with him not having the desire to sound like a native speaker and having strong attachments to his heritage and native language.

5.8 Limitations and Future Directions

While this study allowed an extensive coverage of NAE segments that may be problematic to a number of SHG speakers with ESL, the data is at the same time very limited. All of the data was collected on a single day of M.S.'s life and only a small set of data could be obtained. Nonetheless, I was able to side with Gregg and Chomsky that M.S.'s IL grammar seems to be a homogenous system that does not depend on task or discourse type, and that task type deviations in performance could be attributed to external factors, such as word frequency effects or effects of phonetic environment,

which influenced accuracy scores, but did not reveal that the phonological representation, and therefore the linguistic knowledge M.S has of these NAE sound structures changes throughout task types.

Moreover, for future studies I would include native speaker judgments for tasks exploring suprasegmentals as well as segments, the reason being a very obvious one to me now. While my study explored the production of individual speech sounds, the mispronunciation of one segment alone will usually not hinder comprehensibility or intelligibility, except when crossing phonemic boundaries. Furthermore, L2 speech is produced in spontaneous contexts where it really matters to be comprehensible and understood, and it is usually context that aids the listener to comprehend and understand speech utterances. Thus, even if the speaker produces things like “Dey vent to school togezer”, comprehensibility and intelligibility would likely not be affected. I should add here that, for this very reason, the sentences the native speaker judges listened to were designed in a way that would suppress semantic priming. But generally, I think it is most fruitful to employ native speaker judgments on the suprasegmental level, because when conversing, speakers produce connected speech for which listeners have context available to understand utterances. Thus, on the basis of 9 utterances from connected (read) speech production, it is also difficult to assess M.S.’s proficiency, which is probably most authentically reflected in a semi-spontaneous or spontaneous speech production (Thomson & Derwing, 2015).

Moreover, while this case study was predominantly focused on production of NAE segments, a perception study would have likely provided additional insight into M.S.’s IL grammar and how the sounds of his L2 are stored phonologically. Perceptual

information would allow me to draw stronger conclusions about how M.S.'s production relates his perception of NAE sounds, especially for the phonological contrasts M.S. is producing, and the ones he has not acquired. Perception data could have potentially provided insight into why some of the contrastive segments of English he produced are not distinctive to him, such as NAE /w/ and /v/ or NAE /æ/ and /ɛ/, even though he has instances of accurate production.

Lastly, I had originally intended to include more information on extra-linguistic factors that modern research has concerned itself with lately, which relate to cognition and personality of an L2 learner: motivation, attitude, identity, gender, personality, anxiety, and aptitude. I think these are factors worth considering for a future case study to obtain further qualitative information, and determine how it pertains to L2 production.

5.9 Concluding Remarks

Pertinent to the research questions raised in Chapter 1, this thesis has sought to draw conclusions on the structure of M.S.'s IL grammar by means of analysis of production data obtained from three different task type formalities (wordlist, sentences, and semi-spontaneous production) combined with the analysis of qualitative data obtained about M.S.'s language learning background, and extra-linguistic (psychological and cognitive) factors that shape this IL grammar.

Investigating the composition of M.S.'s IL grammar through production accuracy scores of NAE segments, this thesis found that the formality of the task type does not correspond to the activation of different grammars, as would be suggested by the 'variable competence model' (Tarone 1987; Ellis 1990). In every case, variable

performance on task types for a given segment could be attributed to external (intra-linguistic) factors that influence production performance, which fall under the category of *performance errors*, and are not reflective of M.S.'s linguistic *competence* and mental representations of these NAE sounds of his IL grammar (Chomsky, 1965). This conclusion is supported by the subject's self-assessment of his L2 pronunciation competence, in which he made statements about his perceptual and production ability of NAE segments like /w/ and /v/, which are difficult to contrast for him, interdental in coda position /θ/ and /ð/ which require more effort according to him, as well as /ɹ/, which requires a complex articulatory mechanism. All of these phenomena described by M.S. represented stable phenomena he observed in his L2 pronunciation ability, which in turn suggested their stability in the production data as well.

Considering the amount of time M.S. has been exposed to his L2 English over the time span of 25 years, the analysis of the production data revealed that nativelike pronunciation had been achieved for only one segment, which in turn suggests that for the other segments it will **never** be achieved. This supports the notion that the acquisition of these underlying NAE sounds is fossilized, or has plateaued, and overall represents a homogenous IL. Additionally, the thesis was able to demonstrate that despite the segmental productions being non-nativelike and M.S. employing approximations and overgeneralizations in his speech production, these production errors are not indicative of an "impoverished" L2 grammar. In fact, according to native speaker judges, M.S. is a highly intelligible and comprehensible L2 speaker of English whose IL grammar does not need nativelike form for M.S. to be able to effectively communicate.

In sum, it can be said that the IL grammar – notwithstanding its difference from the L1 system of a native speaker – is a fully developed and complete system representative of M.S.'s underlying competence, which differs from native NAE grammar in *form*, but not in quality or completeness.

Chapter 6 Reference List

- Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: listener perception versus linguistic scrutiny. *Language Learning*, 59(2), 249–306.
- Achirri, K. (2017). Perceiving identity through accent lenses: A case study of a Chinese English speaker's perception of her pronunciation and perceived social identity. MSU Working Papers in *Second Language Studies*, 8, 5–19.
- Archibald, J., & Libben, G. (1995). *Research perspectives on second language acquisition*. Mississauga: Copp Clark.
- Archibald, J. (1998). *Second language phonology*. Amsterdam: John Benjamins.
- Archibald, J., & Young-Scholten, M. (2000). Second language syllable structure. In J. Archibald (Ed.), *Second language acquisition and linguistic theory*. Oxford: Blackwell.
- Baker, A., & Goldstein, S. (2008). *Pronunciation pairs: An introduction to the sounds of English: Teacher's manual* (2nd ed.). Cambridge: New York.
- Barlow, J.A. (2014). Age of acquisition and allophony in Spanish-English bilinguals. *Frontiers of Psychology*, 5, 1–14.
- Bastug, N. (n.d.). *A contrastive analysis of the English and the German sound system*. 1–20. Retrieved June 8, 2015.
- Bauer, L. (2015). Australian and New Zealand English. In M. Reed & J.M. Levis (Eds.), *The handbook of English pronunciation*. Oxford: Wiley-Blackwell.
- Best, C.T., & Tyler, M.D. (2007). Nonnative and second-language speech perception. Commonalities and complementarities. In M. Munro & O.-S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production*. Johns Benjamins.
- Bien, H., Hanulikova, A., Weber, A., & Zwitserlood, P. (2016). A neurophysiological investigation of non-native phoneme perception by Dutch and German listeners. *Frontiers in Psychology*, 56, 1–10.
- Birdsong, D. (1992). Ultimate attainment in SLA. *Language*, 68, 706–755.
- Blevins, J. (1995). The syllable in phonological theory. In J. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 206–244). Oxford: Blackwell.

- Blevins, J. (2004). *Evolutionary phonology – The emergence of sound patterns*. Cambridge: Cambridge University Press.
- Boberg, C. (2008). Regional phonetic differentiation in Standard Canadian English. *Journal of English Linguistics*, 36(2), 129–154.
- Boberg, C. (2015). North American English. In M. Reed & J.M. Levis (Eds.), *The handbook of English pronunciation*. Oxford: Wiley-Blackwell.
- Broersma, M. (2010). Perception of final fricative voicing: Native and nonnative listeners' use of vowel duration. *Journal of the Acoustical Society of America*, 127(3), 1636–1644.
- Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer. Version 6.0.29. <http://www.fon.hum.uva.nl/praat/>
- Bohn, O-S., & Flege, J.E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, 11, 303–328.
- Bohn, O-S., & Flege, J.E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition*, 14, 131–158.
- Bohn, O-S., & Munro, M.J. (2007). *Language experience in second language speech learning*. Amsterdam: J. Benjamins.
- Bongaerts, T., Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the production of foreign language. *Studies in Second Language Acquisition*, 19, 447–465.
- Brockhaus, W. (1995). *Final devoicing in the phonology of German*. Tübingen: Max Niemeyer Verlag.
- Brown, G., Malmkjaer, K., & Williams, J. (1996). *Performance and competence in second language acquisition*. London: Cambridge University Press.
- Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14, 261–290.
- Bylund, E., Hyltenstam, K., & Abrahamsson, N. (2013). Age of acquisition effects or effects of bilingualism in second language in ultimate attainment? In G. Granena & M. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment*. Amsterdam: John Benjamins.
- Carroll, S. (1992). On Cognates. *Second Language Research*, 8(2), 93–119.

- Celce-Murcia, M., Brinton, D., & Goodwin, J.M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.
- Chambers, W.W., & Wilkie, J.R. (1970). *A short history of the German language*. London: Methuen.
- Chamson, E. (2016). Visiting Vienna. *English Today*, 32(4), 40–48.
- Charles-Luce, J. (1985). Word-final devoicing in German: effects of phonetic and sentential contexts. *Journal of Phonetics*, 13, 309–324.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: MIT Press.
- Clements, G.N. (1990). The role of sonority in core syllabification. In J. Kingston, M.E. Beckman (Eds.), *Papers in laboratory phonology I. Between the grammar and physics of speech* (pp 283–333). Cambridge: Cambridge University Press.
- Colantoni, L., & Steele, J. (2006). Native-like attainment in the L2 acquisition of Spanish stop-liquid clusters. In C.A. Klee & T.L. Face (Eds.), *Selected proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as first and second languages* (pp. 59–73). Somerville, MA: Cascadilla Press.
- Colantoni, L., Steele, J., & Escudero, P. (2015). *Second language speech: Theory and practice*. Cambridge: Cambridge University Press.
- Collins, B., & Mees, I. (1984). *The sounds of English and Dutch*. Leiden: Brill.
- Coppieters, R. (1987). Competence differences between native and fluent non-native speakers. *Language*, 63, 544–573.
- Crystal, D. (2011). *A dictionary of linguistics and phonetics* (6th Ed). Oxford: Blackwell.
- Davenport, M., & Hannahs, S.J. (2010). *Introducing phonetics and phonology* (3rd ed). London: Hodder Education.
- Derwing, T.M., & Munro, M.J. (1997). Accent, intelligibility, and comprehensibility. *Studies in Second Language Acquisition*, 20, 1–16.
- Derwing, T.M., & Munro, M.J. (2015). *Pronunciation fundamentals: evidence-based perspectives for L2 teaching and research*. Amsterdam: John Benjamins.
- Derwing, T.M. (2010). Utopian goals for pronunciation teaching. In J.M. Levis, & K. LeVelle (Eds.), *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference* (pp. 24–37). Ams, IA: Iowa State University.

- Derwing, T.M., & Rossiter, M.J. (2002). ESL learners' perceptions of their pronunciation needs and strategies. *System*, 30(2), 155–166.
- Derwing, T. (2003). What do ESL students say about their accents? *Canadian Modern Language Review*, 59(4), 547–567.
- Dickerson, L. (1975). The learner's interlanguage as a system of variable rules. *TESOL Quarterly*, 9, 401–407.
- Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, 62, 284–301.
- Dodd, B., Holm, A., & Hua, Z. (2003). Phonological development: a normative study of British English-speaking children. *Clinical Linguistics and Phonetics*, 17(8), 617–643.
- Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. New York, NY: Routledge.
- Eckman, F.R. (1977). Markedness and the Contrastive Analysis Hypothesis. *Language Learning*, 27, 315–330.
- Eckman, F.R. (1991). The structural conformity hypothesis and the acquisition of consonant clusters in the interlanguage of ESL learners. *Studies in Second Language Acquisition*, 13, 23–41.
- Ellis, R. (1984). *Classroom second language development*. Oxford: Pergamon Press.
- Ellis, R. (1990). *Instructed second language acquisition*. Oxford: Blackwell.
- Ellis, R. (1997). SLA and language pedagogy: An educational perspective. *Studies in Second Language Acquisition*, 19, 69–92.
- Escudero, P., & Wanrooij, K. (2010). The effects of L1 orthography on non-native vowel perception. *Language and Speech*, 53(3), 343–365.
- Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, 36(2), 345–360.
- Flege, J.E. (1987). The production of new and similar phones in a foreign language: Evidence for the effect of the equivalence classification. *Journal of Phonetics*, 15, 47–65.
- Flege, J.E., (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 229–273). Timonium, MD: York Press.

- Flege, J., Schirru, C., & MacKay, I. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication, 40*, 467–491.
- Flege, J.E. (2016). The role of phonetic category formation in second language speech acquisition. *8th International Conference on Second Language Speech*, 10-12 June, 2016, Aarhus University, Denmark.
- Gatbonton, E. (1975). *Systematic variations in second language speech: a sociolinguistic study*. Doctoral Dissertation, McGill University, Montreal.
- Granema, G., & Long, M. (2012). *Sensitive periods, language aptitude, and ultimate L2 attainment*. Amsterdam: John Benjamins.
- Grantham O'Brien, M., & Fagan, S.M. (2016). *German phonetics and phonology: Theory and practice*. New Haven: Yale University Press.
- Gregg, K. (1990). The variable competence model of second language acquisition, and why it isn't. *Applied Linguistics, 11*(4), 364–383.
- Hall, C. (2003). *Modern German pronunciation* (2nd ed.). Manchester: Manchester University Press.
- Hancin-Bhatt, B. (1994). Segment transfer: a consequence of a dynamic system. *Second Language Research, 10*(3), 241–269.
- Hansen Edwards, J. (2008). Social factors and variation in production in L2 phonology. In J. Hansen Edwards & M. Zampini (Eds.), *Phonology and second language acquisition* (pp. 251–283). Amsterdam: John Benjamins.
- Hansen Edwards, J., & Zampini, M. (Eds.). (2008). *Phonology and second language acquisition*. Amsterdam: John Benjamins.
- Hanulikova, A., & Weber, A. (2012). Sink positive: Linguistics experience with th substitutions influences nonnative word recognition. *Attention, Perception, & Psychophysics, 74*(3), 613–629.
- Hardcastle, W.J., & Marchal, A. (Eds.). (2012). *Speech production and speech modeling*. Dordrecht: Kluwer Academic.
- Honeybone, P., & Salmons, J. (Eds.). (2015). *The Oxford handbook of historical phonology*. Oxford: Oxford University Press.
- Hulstijn, J.H. (2015). Explaining phenomena of first and second language acquisition with the constructs of implicit and explicit learning. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages*. Amsterdam: John Benjamins.

- Hyltenstam, K., & Abrahamsson, N. (2003). Maturation constraints in SLA. In C. Doughty & M.H. Long (Eds.), *The handbook of second language acquisition* (pp. 539–588). Oxford: Blackwell.
- James, A.R. (1988). *The acquisition of a second language phonology – A Linguistic theory of developing sound structures*. Tübingen: Gunter Narr Verlag.
- James, A.R., & Leather, J. (1997). *Second language speech: Structure and process*. Berlin: Mouton de Gruyter.
- Jones, D. (1966). *The pronunciation of English*. Cambridge: Cambridge University Press.
- Jordan, G. (2004). *Theory construction in second language acquisition*. Amsterdam: John Benjamins.
- Kennedy, S., & Trofimovich, P. (2008). Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review*, 64(3), 459–489.
- Kent, R.D., & Read, C. (2002). *The acoustic analysis of speech*. Australia: Singular/Thomson Learning.
- Kohler, K. (1999). German. In *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.
- König, E., & Gast, V. (2009). *Understanding English-German contrasts* (2nd ed.). Berlin: Erich Schmidt Verlag.
- Labov, W., Ash, S., & Boberg, C. (2006). *The atlas of North American English*. Berlin: Mouton de Gruyter.
- Ladefoged, P. (2006). *A course in phonetics* (5th ed.). Boston: Thomson & Wadsworth.
- Lado, R. (1957). *Linguistics across cultures*. Tempe: Arizona State University.
- Lenneberg, E.H. (1967). *Biological foundations of language*. Oxford: Wiley-Blackwell.
- Libben, G. (2000). *Representation and processing in the second language lexicon: The homogeneity hypothesis*. Oxford: Oxford University Press.
- Lin, Y-H. (2001). Syllable simplification strategies: a stylistic perspective. *Language Learning*, 51(4), 681–718.
- Lombardi, L. (2003). Second language data and constraints on manner: Explaining substitutions for the English interdental. *Second Language Research*, 19(3), 225–250.

- Maddieson, I. (2013). Presence of uncommon consonants. In M.S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Major, R.C. (1994). Chronological and stylistic aspects of second language acquisition of consonant clusters. *Language Learning*, 44(4), 655–680.
- Major, R.C. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. London: Routledge.
- Major, R.C. (2007). Identifying a foreign accent in an unfamiliar language. *Studies in Second Language Acquisition*, 29, 539–556.
- Major, R.C. (2008). Transfer in second language phonology. In J. Hansen Edwards & M. Zampini (Eds.), *Phonology and second language acquisition* (pp. 63–94). Amsterdam: John Benjamins.
- van Maastricht, L., Krahmer, E., & Swerts, M. (2016). Native speaker perceptions of (non)native prominence patterns: effects of deviance in pitch accent distributions on accentedness, comprehensibility, intelligibility, and nativeness. *Speech Communication*, 83, 21–33.
- McCullough, E.A. (2013). *Acoustic Correlates of Perceived Foreign Accent in Non-Native English*. Doctoral dissertation, The Ohio State University, Columbus.
- Midgley, K.J., Holcomb, P.J., & Grainger, J. (2011). Effects of cognate status on word comprehension in second language learners: an ERP investigation. *Journal of Cognitive Neuroscience*, 23(7), 1634–1637.
- Mielke, J., Baker, A., & Archangeli, D. (2016). Individual-level contact limits phonological complexity: Evidence from bunched and retroflex r. *Language*, 92(1), 101–140.
- Moran, S., McCloy, D., & Wright, R. (Eds.). (2014). PHOIBLE online. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://phoible.org>.
- Moyer, A. (2004). *Age, accent and experience in second language acquisition. An integrated approach to critical period inquiry*. Clevedon: Multilingual Matters.
- Moyer, A. (2013). *Foreign accent: the phenomenon of non-native speech*. Cambridge: Cambridge University Press.
- Munro, M.J., & Derwing, T.M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45, 73–97.
- Munro, M.J., & Derwing, T.M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 49, 285–310.

- Murphy, J.M., (2014). Intelligible, comprehensible, non-native models in ESL/EFL pronunciation teaching. *System*, 42, 258–269.
- Parkes, G., & Cornell, A. (1996). *NTC's dictionary of German false cognates*. New York: McGraw Hill.
- Pätzold, M., & Simpson, A.P. (1997). Acoustic analysis of German vowels in the Kiel Corpus of Read Speech. In A.P. Simpson, K.J. Kohler, & T. Rettstadt (Eds.), *The Kiel Corpus of Read/Spontaneous Speech --- Acoustic data base, processing tools and analysis results* (pp. 215 – 247). Kiel: Institut für Phonetik und Digitale Sprachverarbeitung Universität Kiel.
- R Core Team. (2018). R: A Language and Environment for Statistical Computing. <https://www.r-project.org>
- Richards, J., & Sampson, G. (1974). The study of learner English. In J. Richards (Ed.), *Error analysis: perspectives on second language acquisition*. London: Longman.
- Reed, M., & Levis, J. (2015). *The handbook of English pronunciation*. Oxford: Wiley-Blackwell.
- Richards, J. (1974). *Error analysis: perspectives on second language acquisition*. London: Longman.
- Roettger, T., Winter, B., Grawunder, S., Kirby, J., & Grice, M. (2014). Assessing incomplete neutralization of final devoicing in German. *Journal of Phonetics*, 43, 11–25.
- Rogerson-Revell, P. (2011). *English phonology and pronunciation teaching*. London: Continuum.
- Saito, K., & Brajot, F. (2013). Scrutinizing the role of length of residence and age of acquisition in the interlanguage pronunciation development of English /r/ by late Japanese bilinguals. *Bilingualism: Language and Cognition*, 16, 847–863.
- Salmons, J. (2012). *A history of German*. Oxford: Oxford University Press.
- Sato, C.J. (1985). Task variation in interlanguage phonology. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 181–196). Rowley, MA: Newbury House.
- Seidlhofer, B. (2001). Closing a conceptual gap: The case for a description of English as a lingua franca. *International Journal of Applied Linguistics*, 11(2), 133–158.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, 10, 209–231.
- Simons, G.F., & Fennig, C.D. (Eds.). (2018). *Ethnologue: Languages of the world* (21st ed.). Dallas, TX: SIL International. Online version: <http://www.ethnologue.com>.

- Silveira, R. (2004). *The influence of pronunciation instruction on the perception and production of English word-final consonants*. Unpublished doctoral dissertation, Universidade Federal de Santa Catarina, Florianópolis, Brazil.
- Silveira, R. (2007). The role of task-type and orthography in the production of word-final consonants. *Revista de Estudos da Linguagem*, 15(1), 143–167.
- Skinner, B.F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.
- Smith, B.L., Hayes-Harb, R., Bruss, M., & Harker, A. (2009). Production and perception of voicing and devoicing in similar German and English word pairs by native speakers of German. *Journal of Phonetics*, 37, 257–275.
- Snowling, M.J., & Hulme, C. (2013). Children's reading impairments: From theory to practice. *Japanese Psychological Research*, 55(2), 186–202.
- Steinlen, A. (2005). *The influence of consonants on native and non-native vowel production*. Tübingen: Gunter Narr Verlag.
- Swan, M., & Smith, B. (2001). *Learner English: A teacher's guide to interference and other problems, Vol 1* (2nd ed.) (S. Thornbury, Ed.). Cambridge: Cambridge University Press.
- Szubko-Sitarek, W. (2014). *Multilingual lexical recognition in the mental lexicon of third language users* (2015 ed.). New York: Springer.
- Tarone, E. (1987). The phonology of interlanguage. In G. Ioup & S. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 70–85). Rowley, MA: Newbury House.
- Tarone, E., & Parrish, B. (1988). Task-related variation in interlanguage: The case of articles. *Language Learning*, 38, 21–44.
- Thomson, R.I., & Derwing, T.M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*, 36(3), 326–344.
- Ulbrich, C., & Ordin, M. (2014). Can L2-English influence L1-German? The case of post-vocalic /r/. *Journal of Phonetics*, 45, 26–42.
- Watkins, M.A., Rauber, A.S. & Baptista, B.O. (2009). *Recent research in second language phonetics/phonology: perception and production*. Newcastle upon Tyne: Cambridge Scholars.
- Wells, J.C. (1982). *Accents of English I: An introduction*. Cambridge: Cambridge University Press.

- White, L., & Genesee, F. (1996). How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research*, 12(3), 233–265.
- Whitley, M.S. (2002). *Spanish/English contrasts*. Washington, DC: Georgetown University Press.
- Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag: New York.
- Wode, H. (1980). Phonology in L2 acquisition. In S. Felix (Ed.), *Second language development*. Tübingen: Gunter Narr Verlag.
- Wode, H. (1981). *Learning a second language. An integrated view of language acquisition*. Tübingen: Gunter Narr Verlag.
- Yang, C. (2016). *The acquisition of L2 Mandarin prosody: from experimental studies to pedagogical practice*. Amsterdam: John Benjamins.
- Yildiz, Y. (2010). *Age effects in the acquisition of English onset clusters by Turkish learners: An optimality theoretic approach*. Newcastle upon Tyne: Cambridge Scholars.
- Yule, G., & Tarone, E. (1991). The other side of the page: Integrating the study of communication strategies and negotiated input in SLA. In R. Phillipson, E. Kellerman, L. Selinker, M. Sharwood Smith, & M. Swain (Eds.), *Foreign/Second language pedagogy research* (pp. 162–171). Clevedon: Multilingual Matters.
- Zec, D. (1995). Sonority constraints on syllable structure. *Phonology*, 12(1), 85–129.

Chapter 7 Appendix

Appendix A

Background Questionnaire (Lin, 2016)

Name: _____ Age: _____
Gender: _____ E-mail address: _____
Place of birth: _____ Profession: _____

1. What is the highest level of education you have completed? Please list certificates or degrees that you received.
2. Do you have any specific linguistic qualifications? For example: are you a translator, interpreter, a foreign language teacher, or anything of that kind?
3. What language did you learn first?
4. What language did you grow up speaking, if different from (3)?
5. Where did you grow up, if different from your place of birth?
6. What was the first foreign language that you learned, and at what age did you learn it?
7. At what age did you have your first contact with English?
8. What major variety or dialect of English are you most familiar with (i.e. British English, Canadian English, etc.)?
9. For how long have you been learning English?
10. What countries have you stayed or lived in where English is spoken, and how did you stay in each one of them?

11. Please list the languages that you have learned, with the age that each was learned and indicate how well you can speak these languages now (by giving a number from 1 to 5 where 1 means “poor” and 5 means “native-like”).

Language	Age	Level
		speaking___, listening___, reading___, writing___
		speaking___, listening___, reading___, writing___
		speaking___, listening___, reading___, writing___
		speaking___, listening___, reading___, writing___

12. For how long have you been in Canada?
13. How many hours per day do you speak with people who are fluent English speakers?
14. Do you have any hearing loss, or any other kind of impairment affecting your hearing?
15. Have you taken any of the following proficiency tests? If not, please mark “N/A”, and, if yes, please provide scores and the year you took the test.

Test	Score	Year
TOEFL (Test of English as a Foreign Language)		
MELAB (Michigan English Language Assessment Battery)		
IELTS (International English Language Testing System)		
Other		

Appendix B
Production Wordlist Task Tokens

Each slide presented one token at a time to the participants M.S. and C.E. in a PowerPoint Presentation:



cup

weather

picture

Appendix C

The sentences for the production task were also presented on a computer screen in a PowerPoint presentation.

Instructions:

1. Please read out the following sentences. 2. Don't alter your natural pace or intonation.

- A beautiful tree I see, but a printer code can't copy it.
- The lab partner evolved from the wallpaper the bad man had left.
- Sarah takes out her contacts because her ear drum has had enough of the zigzag music.

Appendix D

In the second part of the interview served as a questionnaire and self-assessment of English pronunciation for M.S. Additionally, questions pertaining to extra-linguistic factors were asked.

1. How do you like life in Canada?
2. What brought you to Canada?
3. What aspects of Canadian culture appeal to you the most?
4. Do you miss Germany sometimes?
5. What aspects of German culture appeal to you the most?
6. Would you say “being German” is important to you?
7. Learning English in school, did your teachers teach you much about pronunciation? If yes, what pronunciation rules do you recall?
8. How important is it to you to have good English pronunciation?
9. Are there any sounds you find difficult to pronounce?
10. What do you like the most about the English language?

Appendix E

This questionnaire served as an addition to the transcriptions and goodness ratings the native speaker judges provided for the intelligibility, comprehensibility, and accentedness measurements of M.S.'s and C.E.'s speech production (partially adapted from Murphy, 2014)

Provide comments, if you think you can:

1. I recognize the accent I have just listened to.
2. I noticed some particular things about some individual sounds.
3. There were very few segmental errors.
4. Segmental errors made the speaker difficult to understand.
5. One speaker was easier to understand than the other.